Journal of the
OPTICAL SOCIETY *of* AMERICA

VOLUME 35, NUMBER 8 AUGUST, 1945

Schlieren and Shadowgraph Equipment for Air Flow Analysis

NORMAN F. BARNES AND S. LAWRENCE BELLINGER
General Engineering and Consulting Laboratory, General Electric Company, Schenectady, New York
(Received February 15, 1945)

THOUGH the science of making the flow of air visible to the human eye is still in its infancy, the fundamental principle dates back about 86 years. At that time Foucault devised a test, which now bears his name, for observing the surface configuration of a concave telescope mirror. Though in his tests Foucault probably observed what were apparently heat waves passing in front of his mirror surface, it remained for Toepler to realize the real significance of this additional phenomenon. Since the latter's publication in 1864, the method has been often referred to as the Toepler Schlieren Method.

The operation of the schlieren method can be best described with reference to Fig. 1. Light from an illuminated pinhole S is allowed to fall upon lens L_1 and be converged to the image point at P. If the eye is placed slightly behind the image point, as in the Foucault test, the lens L_1 will appear as a uniformly illuminated field. In usual practice, however, a lens L_2 is substituted for the eye, this lens focusing the striation upon a viewing screen or upon a photographic plate. If the knife-edge E is then moved laterally across the image point until all the rays passing through that image are obscured, the field as viewed upon the screen will be uniformly dark.

Next to be considered is the effect of introducing air into a very small element of volume dv (Fig. 1). The index of refraction n of this air is different from that of the surrounding air n_0 for which the knife-edge has been adjusted to produce a uniformly dark field upon the viewing screen. A light ray passing through the volume dv will be refracted. Therefore, it will no longer proceed along its original path, except for the singular instance for which the gradient or change in the index of refraction as a function of the space position is parallel to that light ray.

If the gradient of n through the refraction is so directionally constituted that the light ray is refracted in a counterclockwise manner, this ray

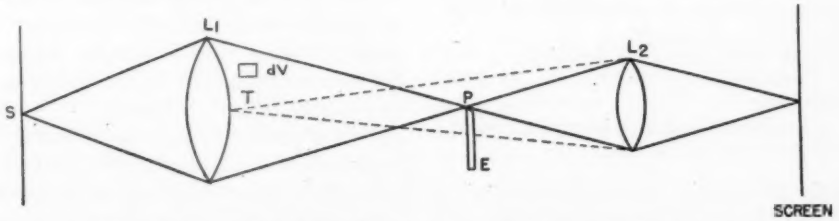


FIG. 1. Schematic diagram of the basic elements of a lens schlieren system.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

will no longer pass through the image point P but will travel above it. Hence, this ray will not be obscured by the knife-edge but will pass on to the screen. There it will illuminate a point corresponding to the location of the elementary volume dv. Thus for every point in the region T for which a similar refraction takes place, there will be a corresponding point illuminated on the viewing screen. The composite of all such points forms the image of the phenomenon being investigated.

If the refraction in the elementary volume dv is in a clockwise direction, the rays will be bent down toward the base of the knife-edge, and the corresponding points on the viewing screen will be dark.

In making the lateral adjustment of the knife-edge when no air flow or striation is present in the region T, it is usually desirable to allow some of the rays to pass over the knife-edge in order to produce a uniformly low background illumination. The presence of this background makes it possible to see more clearly in silhouette or outline the objects or models used in producing the airflow phenomena. Thus, referring to the earlier paragraphs about the refraction produced by the elementary volume dv, a counterclockwise bending of a light ray will produce an increase in illumination at the corresponding screen point; a clockwise bending will produce a decrease in screen illumination.

It is apparent from the foregoing that, if the knife-edge is allowed to cut the light beam laterally from the top rather than from the bottom, the resultant screen pattern will be the negative of the former one. In other words, where the picture was previously light, that same area will now be dark.

The refraction of the light rays in the region T can be produced in four ways:

(a) by a change in temperature of the fluid being investigated—air, for example;
(b) by a change in the pressure of the fluid;
(c) by the introduction of a transparent object into the light path; and
(d) by the use of a gas having an index of refraction different from that of air.

Since this method is extremely sensitive to light refraction produced either as a result of a continuous index of refraction gradient or as a result of surface refraction, it is often useful for checking the optical homogeneity and surface flatness of transparent materials, such as glass.

For an effectively circular light source, such as an illuminated pinhole, the rays which contribute to the greatest sensitivity of the system pass through the region of the light-source image near the knife-edge. Thus rays initially passing through this point have to be refracted upward only slightly in order for them to pass over the knife-edge and contribute to the illumination on the viewing screen. On the other hand, those rays which pass through the lower part of the light-source image away from the knife-edge must undergo a large refraction in the region T of Fig. 1 before they can contribute to the screen illumination. For the highest sensitivity, therefore, the maximum number of rays in the light-source image must lie close to the knife-edge. This requirement is most easily satisfied by the use of a slit source of light. Here, no rays lie at a large distance from the edge itself; consequently, more of the rays can be bent over the knife-edge for a given amount of refraction in the striation.

An additional requirement for maximum sensitivity is that the light-source image adjacent to the knife-edge be as straight and as sharply defined as possible. Then it will conform to the sharp edge of the knife-edge so that accurate positioning of the knife-edge can be made with respect to the image.

If the index of refraction gradient lies only in a horizontal plane, thereby producing a corresponding refraction in that plane, the effect is to produce a lateral shift of an affected ray parallel to the knife-edge. Such a displacement, if strictly parallel to the edge, cannot affect the viewing-screen illumination. It can be seen, therefore, that a further requirement for maximum sensitivity of the system is that the effective light source, its image, and the knife-edge must be perpendicular to the index of refraction gradient being considered. An optical method for adjusting the light source and image orientation will be described later.

This alignment requirement need not be absolutely adhered to. Fundamentally, the foregoing statements are true only if all the rays incident upon the air flow being investigated are mathematically parallel, and if there is no index of

refraction gradient component at right angles to the parallel-ray direction. Both of these conditions are seldom, if ever, completely realized because of the optical aberrations of the system and of the complexity of the refraction gradient. This is particularly true for the lens-type system shown in Fig. 1. Here, the refraction occurs in the region T normally traversed by convergent rays. For systems described later, in which effectively parallel light is used in the striation region, the requirement of correct knife-edge orientation with respect to the gradient of n is more important. However, even here the introduction of coma and astigmatism into the refracted rays is often sufficient to make a strong gradient visible even though its direction is parallel to the light source and the knife-edge.

Thus far it has been shown that the effective light source in the form of a slit will allow much greater sensitivity than an effectively circular light source and that the source should be bordered, at least on one side, by a straight, sharp edge. The latter requirement can be satisfactorily met even though it may be necessary to use an irregularly shaped light source, such as a coiled tungsten filament. If this is so, an auxiliary condenser system must be used to project an image of the filament away from the lamp so that a knife-edge, slit, or other aperture can sharply limit the geometrical shape or size of the transmitted image. Hence, there is produced an effective light source that is, in the present example, sharply bounded on one side by a straight edge.

The relative aperture of the condenser system must be at least as large as that of the schlieren system itself. Otherwise, the striation field will not be fully illuminated. Furthermore, if the light source is large or extended laterally so that it cannot be considered as having its entire area effectively located on the optical axis of the system, then the aperture of the condenser system must be larger than that of the schlieren system so that the cone of rays emerging from each point of the source contributes to the general illumination in the entire striation field. Otherwise, light rays coming from those points of the light source farthest from the optical axis will not spread out over the entire striation field but will be vignetted by the condenser aperture.

Uniform surface luminosity of the effective light source is also necessary for maximum sensitivity and for uniformity of the striation field.

A circular light source can be successfully used if the corresponding knife-edge has the form of a circular disk of such a size that its edge coincides with that of the light-source image itself. The advantage of a system using this type of source and knife-edge is that it is uniformly sensitive to the index of refraction gradients which lie in any direction in a plane perpendicular to the rays incident upon the striation field. If maximum sensitivity is not required, a straight knife-edge can be used with a circular source, the knife-edge being rotated about the optical axis until its position is perpendicular to the direction of the index of refraction gradient of major interest.

Instead of rotating the knife-edge, as just described, two straight knife-edges can be used in the same plane and at right angles to each other. They simultaneously cut the light-source image from two sides, vertically and horizontally, for example. Thus, the index of refraction gradients will have projected components perpendicular to at least one of the two knife-edges.

If it is desirable to use an irregularly shaped light source, such as a coiled tungsten filament, without the use of a condenser system and auxiliary knife-edge, as previously described, an arrangement known as the masking method can be used. In this system the optical components are aligned as before, with the exception that no knife-edge is placed in the plane of the light source image. Instead, a photographic plate is inserted in that plane, and an exposure to the light source image is made upon the plate. After processing, the plate will be the negative of the image, and, if replaced in the image plane in the proper position, will completely mask the light converging to that image. As in the case of the circular light source and the circular knife-edge disk, this system is sensitive to index of refraction gradients lying in any direction in the plane perpendicular to the rays passing through the striation. From the standpoint of physical stability of the image, sharpness of the image, and maximum light transmission through the plate background surrounding the image, lantern slide plates will be found to be very satisfactory.

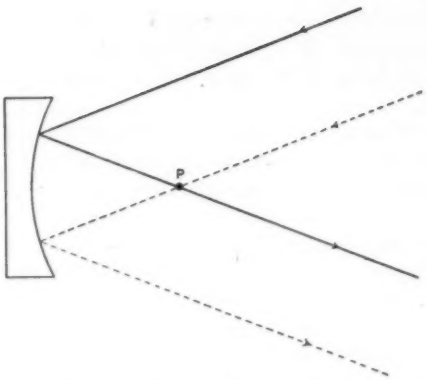


FIG. 2. Double passage of light through a point in the striation region.

It must be pointed out, however, that with irregularly shaped sources this method cannot be satisfactorily applied to quantitative analysis problems or to problems in which good interpretation of the initial index of refraction gradients is necessary.

Referring again to Fig. 1, the striae which can be studied using a single field-lens system are those which effectively occur in a plane, since the light passing through the striae is convergent. If the striae are 3-dimensional rather than effectively 2-dimensional, the results will be much more satisfactory if the light incident upon them is not convergent but parallel, as can be the case if two field lenses are used.

The field lenses used in the schlieren systems must be well corrected lenses, particularly from the standpoint of spherical aberration. Otherwise, it will not be possible to obtain a uniform brightness across the field projected upon the viewing screen. Also, if large amounts of chromatic aberration are present, the striation image on the viewing screen will not be sharp. The glass of the lenses must be of the finest optical quality and free from scratches so that the lenses will be striation free. Otherwise, any striae in the field lenses will be superimposed upon those which are being investigated.

If a large diameter field is required, it is exceedingly difficult, if not practically impossible, to obtain well-corrected lenses. For that reason it is highly desirable to use concave mirrors in place of the lenses. The use of such mirrors has many advantages. Since first surface mirrors are used, the optical quality of the glass does not affect the striation field, provided the surface is free from air bubble pockets. Furthermore, chromatic aberration is eliminated. Since the mirror surface can be parabolized, it is possible to obtain large mirrors which are well corrected for spherical aberration.

A single mirror system has the disadvantage that light passes through a point in the striation twice. In Fig. 2, a ray of light will pass through point P before it strikes the mirror surface, and another ray will pass through P after leaving the mirror surface. Thus, each ray will pass through two different points of the striation so that it will undergo a double refraction. If the striation is very weak and if the angle between the incident and reflected rays is small, this double refraction will give the system twice the sensitivity that it would have if the rays were to pass through the striation once. For strong striae, however, the doubling of the image and the double refraction of the rays make the resultant image on the viewing screen practically impossible to interpret.

Figure 3 shows a system of high sensitivity for the investigation of very weak striations, viewed in parallel light. Using the same nomenclature as before, S is the effective light source, P, the source image, and T, the striation region. The axis of the plane mirror M_2 is so adjusted that it makes a very small angle with the axis of the concave mirror M_1. As a result, the image of the source and the source itself are separated so that the corresponding apparatus parts do not have to occupy the same physical space. For minimum and almost negligible doubling it is necessary to keep the striation region as close

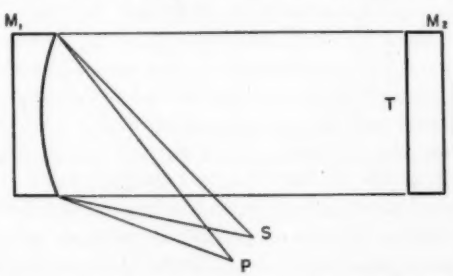


FIG. 3. System of high sensitivity for investigation of very weak striation viewed in parallel light.

to the plane mirror as possible. In the region common to the light cones subtended from S and P and the parallel beam between the mirrors, there will be a very large amount of doubling.

For stronger striae, the system shown in Fig. 4 has proved to be the most satisfactory. Here two concave mirrors are used, and the light passes through the striation region only once. Since this system is free from image and refraction doubling, it will allow the highest resolving power, and though the theoretical sensitivity is only half of that of the system shown in Fig. 3, it has proved to be ample. Any loss of sensitivity has been more than offset by the increase in resolving power. If the light source and its image as formed by the second mirror lie on opposite sides of the common mirror axis and in the same plane as shown by the solid lines, the system is optically symmetrical and effectively coma free. If the light source image lies on the same side as the light source itself, as shown by the dotted lines, bad coma is present, and satisfactory results cannot be obtained. In both Figs. 3 and 4, the off-axis angles are exaggerated. In actual practice the concave mirror axes must be slightly displaced from the common mirror axis in order to remove the light source and its image from the parallel beam.

In installation, schlieren systems can, as has already been shown, be set up in any of a number of different ways, so the type of system best suited to the particular needs can be chosen. In the twin-mirror system being described, six types of arrangement can be employed. For all six arrangements, the position of the two mirrors can remain practically unchanged. Variations are accomplished by the repositioning of the light source and the knife-edge assemblies relative to the common mirror axis.

The basic arrangement is that in which the slit source, mirror axes, and knife-edge are all in the same horizontal plane. If the plane of major refraction gradient is vertical, the plane containing the center of the slit source, the mirror axis, and the center of the knife-edge will be horizontal. If the gradient is in a horizontal plane, this arrangement can be changed so that the slit source and knife-edge plane is vertical.

The next two arrangements are very similar, except that the plane of the entire system is vertical instead of horizontal. Here again the light source and knife-edge may be either horizontal or vertical, depending upon the direction of the refraction gradient being studied.

The foregoing four arrangements are to be preferred for reasons to be discussed later, and it will be shown that two of these will prove to be much better from an optical point of view than the other two.

The final two arrangements are those in which

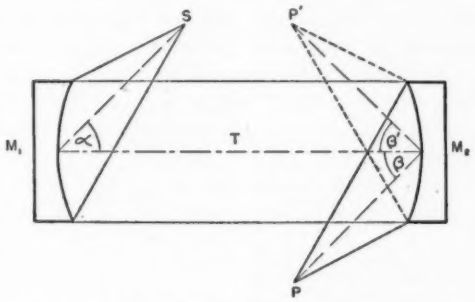


FIG. 4. Twin mirror system for obtaining maximum resolving power.

the light-source center and the common mirror axis are in a vertical plane, and the knife-edge center and the mirror axis are in a horizontal plane, or *vice versa*. In this arrangement, the light source and knife-edge can again be either horizontal or vertical, as dictated by the conditions being studied. This system presents a number of serious optical difficulties which will be treated in the section dealing with aberrations involved in the different types of arrangements. It will be apparent that the all-horizontal and all-vertical systems outlined above are optically identical, except that the whole system is rotated 90 degrees around the common mirror axis when the set-up is changed from one to the other.

When setting up the twin-mirror system, as shown in Fig. 4, it is essential, first, that angle α be as nearly equal to angle β as possible; and second, that these angles be kept as small as possible. The equality of these angles is required by the necessity of keeping the system symmetrical in order to avoid coma, while the size of the angles will determine the amount of astigmatism which is introduced. Either of these

aberrations will cause a poor knife-edge image of the light source, with a resultant loss in the uniformity of the light cut off by the knife-edge. This in turn results in uneven sensitivity over the schlieren field. It is also most important that all the components of the system be in the same plane since, if this were not so, coma would again be introduced. Thus, the ideal condition requires that the whole system be symmetrical and in one plane.

As has already been mentioned, the shapes of the light-source image and the knife-edge must conform to each other in order to obtain uniform sensitivity. Where a slit source and straight knife-edge are employed, as in the system described, it is important, therefore, that the system be so arranged that the astigmatism shall not cause a deformation of the light-source image in a plane perpendicular to the plane of the knife-edge. It has been found that, if the system is so arranged that the astigmatism merely lengthens the light-source image at the knife-edge, but does not broaden it, the sensitivity and light distribution over the schlieren field will not be detrimentally affected. From this it now becomes apparent that, if a horizontal striation is to be studied, the most ideal arrangement of the system will be to have all components in a horizontal plane, with the light source and knife-edge horizontal as well.

Similarly, a study of vertical striations can best be carried out with the entire system rotated 90 degrees about the common mirror axis. The reason for this is that, with the knife-edge and light source both in a horizontal position, the astigmatic aberration is required to be in a horizontal direction also, in order to lengthen rather than broaden the light-source image at the knife-edge. The same applies to the use of a vertical slit source and knife-edge.

In further consideration of this point, it can now be seen that the system described previously, in which the light source and knife-edge are in planes at 90 degrees to each other relative to the common mirror axis, is most disadvantageous for several reasons. Since the combination of the light source and the first mirror form an off-axis system, this combination becomes a source of coma. By the same token, the second mirror and the knife-edge form another off-axis system, but in a different direction; so this combination is a further source of coma.

The effect of coma from two different portions of the system cannot be cancelled unless the system is symmetrical and all in the same plane. Therefore, the light-source image at the knife-edge is distorted by the presence of two comatic images. But this is not all. In addition, there is astigmatism to consider. One half of the system will introduce astigmatic aberration which will elongate the light-source image. It has already been seen that this is not a serious detriment. However, the other half of the system will introduce astigmatic aberration which will broaden the light-source image at the knife-edge. Hence, it has been found that with this type of optical arrangement, the aberrations are so serious as to make successful results most difficult, if not impossible, to obtain.

For its light source, the schlieren system being described employs a Mazda Type B-H6 lamp. This is a 1000-watt, high pressure mercury lamp used with air cooling. The more familiar Type A-H6 water-cooled lamp was originally tried. It was found, though, that multiple optical difficulties resulted from bubbles forming inside the cooling jacket, and also from the presence of the two additional cylinders of glass in the jacket through which the light had to pass. The air-cooled lamp requires no jackets at all. Therefore, it eliminates these and other problems because it presents to the optical system a sharply defined, clean cut light source. By the nature of the lamp itself, this light source is a natural slit source whose dimensions are about one by twenty-five millimeters.

In instances where it is desired to project upon a viewing screen an enlarged image of the flow model being studied, the full length and breadth of the lamp can be used as a source. This makes the knife-edge image of the lamp quite large, but very good schlieren is obtained nonetheless. With this large source, the chief detriment is a slight loss of image resolution through aberrations resulting from the lateral extension of the ends of the lamp from the axis of the cone of light illuminating the first mirror. This is only a minor difficulty, however, and does not seriously interfere with the use of the system for viewing

magnified images of flow conditions in a steady or slowly changing state.

For aligning the optical system and focusing the camera, the lamp is operated in the normal fashion, but a very small slit aperture is placed close to the lamp. This aperture approximates a point source since it is a slit only ten thousandths of an inch wide and forty thousandths of an inch long.

Photographs may be made with the lamp running continuously, although, to render a good photographic exposure, the use of a fairly heavy absorption filter is necessary in the optical system in order to reduce sufficiently the quantity of light reaching the film. With this combination, photographs may be made on standard photographic materials with a shutter setting of 1/200 of a second. If an air jet, or other high velocity steady-state condition is being studied, it can be readily seen that a photographic exposure of this duration will have an averaging effect, and will, in effect, be a time exposure. In other words, if the light intensity were reduced sufficiently, and a picture taken with a ten-second exposure instead of 1/200 of a second, the resultant pictures would look just about the same.

If there is a high frequency oscillation involved in the jet (as is frequently the case), then such a photograph will lack detail and will not prove to be satisfactory. To secure effectively instantaneous photographs of conditions such as these, an adaptation of the circuit from the General Electric Photolight has been included in the schlieren light-source power supply. This circuit makes it possible to use the B-H6 lamp as a very short-duration flash tube as well as for a continuous light source. This situation offers the very great advantage that it places at the operator's disposal a continuously burning lamp for adjustments and visual studies, and a short-duration flashing lamp for very short exposure photographs, all in one unit. This consequently results in a great saving of time since it does not involve the effort and readjustment required if separate types of light sources are used for continuous and flashing service, respectively. With the system described here, all that is required when changing the lamp from burning to flashing operation is to throw a switch. The system is then ready for instant use in the new operation, the same lamp being used for both continuous burning and flashing.

Photographs made with the photolight circuit have an exposure duration of about four microseconds or less. They are made by opening the camera shutter, flashing the lamp, and closing the shutter again, or by synchronizing the flash with the shutter operation. The extremely short duration of the illuminating flash makes it possible to study very rapidly changing transient conditions, and to secure well-defined photographs of conditions of turbulence and shock within a jet which is operating at considerably higher than sound velocity.

As has been mentioned previously, the sensitivity of the system may be varied considerably by changing the amount of knife-edge cut-off at the light-source image. If the maximum sensitivity is desired, it has been found that the best setting is one in which the field background is slightly illuminated. This makes it possible to see very easily the decreases as well as increases in field illumination which result from refraction gradients in the schlieren field. If the system is used with the background very dark, then only the increases in field illumination can be seen, and the resultant picture contains less detail.

At the setting for maximum sensitivity, the system shows very strongly the warm air which rises from a human hand held in the field. Also, the system will show very nicely a temperature change of about 1°C in air, which corresponds to a change of about one part in a million in its index of refraction. A reduction in sensitivity may also be accomplished by decreasing the amount of cut-off at the knife-edge so that the field illumination is considerably brighter than at the point of maximum sensitivity. The manner in which the sensitivity is adjusted is dictated by the nature of the effects under consideration, and is very simply accomplished by the operator.

The two concave mirrors used in the system shown in Fig. 4 are a matched pair having a diameter of 12 inches and a focal length of 96 inches. They are made of Pyrex glass and are about two inches thick. The concave first surface of each mirror has been aluminized. These mirrors are parabolized to better than $\frac{1}{16}$

wave-length of sodium light.[1] It is absolutely essential from the standpoint of maximum sensitivity and resolving power as well as uniformity of field that the mirrors be of the finest astronomical quality.

As may well be imagined, in a system of this type the nicety of adjustments involved requires special consideration in the construction of the component parts. Of special importance in this respect is the physical stability of all components in relation to each other. The lamp house, mirror mounts, knife-edge, and camera supports must all be sturdily built and so constructed as to eliminate vibration and lack of integration, or at least to reduce them to an absolute minimum. Also, the adjustability of all components, in each of several planes, must be sensitive and accurate.

As a result of these considerations, the mirror-mount supporting stands for this system were built to weigh about half a ton apiece. Each one is mounted on three castered wheels for mobility. When the position of the mirror has been set, screw jacks are lowered to the floor and the castered wheels raised. Fine leveling and vertical adjustments are then accomplished by the manipulation of the three screw jacks on each stand. The mirror mount itself utilizes a three-point suspension of the mirror. One of these points is a universal joint, carefully built to eliminate backlash in its operation. The other two points are closely fitted ball-and-socket joints on large but finely threaded screws. By the manipulation of these screws, very precise control over the alignment of the mirror is made possible. A picture of the schlieren equipment is shown in Fig. 5. The arrangement is that of Fig. 4.

The stands which support the light-source and knife-edge assemblies are somewhat lighter in construction, but are equally sturdy. Jacks and casters are also provided on these. In addition, there are no less than five different adjustments in the knife-edge assembly alone, all of which contribute to adequate control of the system. These five adjustments are:

(1) Fore and aft adjustment, to get the knife-edge exactly at the focal point of the light-source image;

[1] These mirrors were made by the Perkin Elmer Corporation, Glenbrook, Connecticut.

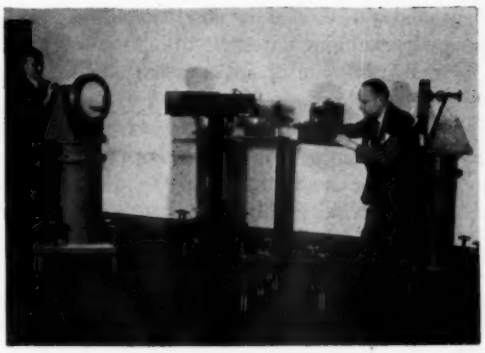


FIG. 5. Picture of the schlieren equipment, showing the light-source unit, knife-edge, camera assembly, and twin parabolic mirrors. The set-up is the same as that in Fig. 4.

(2) Fine vertical adjustment, for control of the amount of cut-off when a horizontal knife-edge is being used;
(3) Rotation of the knife-edge in a horizontal plane, to accommodate the knife-edge to slight lateral variations in focal length;
(4) Rotation of the knife-edge in a vertical plane, to position the edge for either vertical or horizontal light-source image positions;
(5) Cross feed, for adjustment of the amount of cut-off when a vertical knife-edge is being used.

The mounting of the B-H6 lamp in the lamp-house assembly is quite simple. The two brass ferrules at the opposite ends of the lamp are held in brass electrodes which are connected to the power supply. The two air-cooling jets blow across the lamp horizontally, and the air is carried off through a baffle which contains a light trap. The lamp itself is in a horizontal position. In the top of the lamp house there is an aperture through which the light passes to reach the remainder of the optical system.

The adjustment of the position of the slit source relative to the system axis is accomplished by the use of a prism, which is mounted on the top of the lamp house. If an apparently vertical slit is desired, the lamp house is turned so that the axis of the horizontal lamp points in the direction of the first mirror. The prism is then adjusted so that it directs the light coming up from the lamp outward to the surface of the first mirror. The resultant effective position of the source is vertical, and its effective height is at the center of the prism's hypotenuse. It is necessary, therefore, that the lamp-house stand

be so adjusted that the center of the hypotenuse is in the plane which includes the knife-edge center and the common mirror axis.

If a horizontal slit source is desired, the lamp house is so rotated in a horizontal plane that the axis of the lamp is perpendicular to the axis of the cone of light through which it illuminates the first mirror. The prism is again adjusted so that it directs the light to the first mirror; thus an effectively horizontal light source is obtained. This type of manipulation is necessary because the H6 lamps cannot be burned in a vertical position. If this is done, all of the mercury tends to collect in the bottom end of the tube, and the electrode in the upper end disintegrates. (This statement applies only to a.c. operation of the lamp. If d.c. is used, the lamp may be burned vertically provided the lower electrode is always positive.)

The effective distance of the light source from the first mirror is very critical. As has already been pointed out, in order to have the beam through the field made up of light as nearly parallel as possible with this type of system, the effective distance of the light source from the first mirror must be exactly the focal length of the mirror. This adjustment is obtained by slight fore and aft movement of the lamp house until the light which falls on the second mirror just reaches its outer edges all the way around. Since the two mirrors are of the same diameter, this condition shows that the source is at the focal point of the first mirror. After this has been done, the knife-edge is then positioned at the image focus at the other end of the system.

An investigation has been made of the effect upon the resolving power of the system produced as a result of using various distances between the first mirror, the object, and the second mirror. From the standpoint of classical theory, an object placed in any part of the parallel-light region between the two mirrors will be in focus upon the viewing screen at all times. However, when refraction of the rays takes place in the striae, these rays are no longer parallel to the optical axis between the two mirrors, and they must be considered as skew rays. For this reason the striation region must be placed at a distance greater than the focal length of the second mirror away from that mirror. Otherwise the projector lens or camera lens will not resolve a sharply defined image of the striation.

Theoretically, there should be no upper limit to the possible distance between the first and second mirrors. However, there is a very definite lower limit. It has been shown that an object in the region where the parallel beam and either cone of light overlap will produce a double image in the optical system. It is immediately apparent, therefore, that the object must be sufficiently far from the first mirror to avoid any interference in the cone of light which illuminates that mirror, and, as just shown, it must be farther from the second mirror than the knife-edge is.

In practice, using this system in which the matched mirrors are of eight-foot focal length, a spacing of eighteen feet between mirrors has been found to be convenient for most applications. However, there is one factor which may make it desirable at times to try to work with a somewhat shorter distance. That factor is the background disturbance which will appear in the picture as a result of the circulation of normal air currents around the room.

It must be remembered that any and all disturbances of the air in the entire region between the light source and the knife-edge will produce their own schlieren images on the screen, especially where the system is being used at peak sensitivity. If maximum sensitivity must be used, it then becomes important to control, and to reduce to an absolute minimum, turbulent air conditions in the vicinity of the apparatus. Often this will mean that, once the set-up is completed and ready to be photographed, all ventilators must be stopped, doors closed, and the operator seated quietly at his control box for several minutes before actually taking the picture. However, where strong disturbances, such as high pressure jets are being studied, the effect of background can be practically eliminated by somewhat increasing the amount of cut-off at the knife-edge, and thus reducing the sensitivity of the system.

If one places a viewing screen between the jet, or other object being studied, and the second mirror, the image of the jet as seen on the screen will show what is commonly referred to as a shadowgraph. Its formation is also due to changes in the refraction of light by conditions within

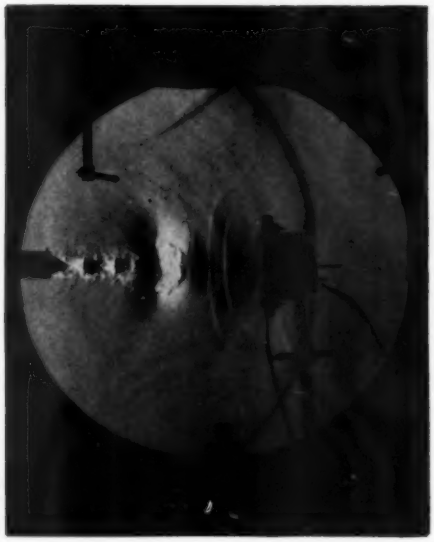


FIG. 6. Schlieren photograph of a jet, showing a sound wave and its reflection from the flash triggering microphone.

the jet. However, there is a distinct difference between a shadowgraph image and a schlieren image.

A shadowgraph shows only the boundaries of sharp wave fronts in the medium, while a schlieren picture shows small gradients as well as wave fronts. If the viewing screen is moved up close to the object, the shadowgraph practically disappears. As the screen is moved farther and farther away, the image appears with ever-increasing contrast upon the screen. The light, as it passes through the boundary layers of the sharp wave fronts, is subjected to refraction very much as it would be if it were to pass through the edge of a very thin-walled glass cylinder, at an angle perpendicular to the cylinder axis and tangent to the cylinder. As it goes through, it is slightly deflected, so that, directly beyond the cylinder, the image shows a decrease in light intensity immediately adjacent to an increase in light intensity. The longer the optical lever arm, the wider this contrasting band becomes, and accordingly, the greater the sensitivity of the system. However, a point is soon reached beyond which the resolution of the image rapidly deteriorates. Thus, a compromise position is chosen in the making of shadowgraphs. The position is generally the point nearest to the object which will give adequate sensitivity.

While work of this type may be carried out in the parallel beam between the two schlieren mirrors, it can be accomplished much more simply and with somewhat better results in another way. All that is required is the object, a spark-gap unit, and a film holder. No lenses are involved. When making the shadowgraphs, some experimentation was required in order to arrive at the distance ratios which would give the most satisfactory results. In most instances, the spark gap was placed about fifteen feet from the jet on one side of, and perpendicular to its axis, while the film was held at a distance of eighteen inches to the other side of, and parallel to, the jet axis.

There is an important reason for using a spark gap in the place of the H6 lamp; namely, circuit characteristics of present equipment make possible a considerably shorter exposure with the gap than is possible with the lamp. The duration of the spark gap exposure is about one-fifth of a microsecond, while the duration of the lamp's flash is about twenty times as long. If the lamp were masked down to the size of the spark gap, there would be ample light to secure a good shadowgraphic exposure, but it has developed in practice that shadowgraphs made with the spark gap produce more favorable results.

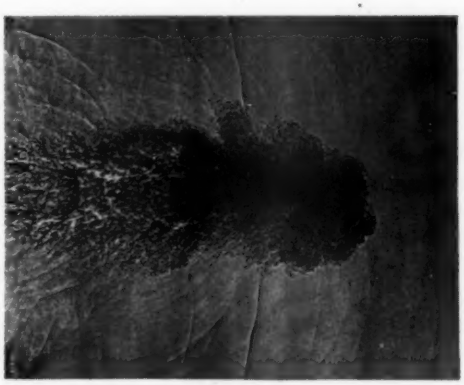


FIG. 7. Shadowgraph of a bullet being discharged from a gun. The many curved lines in the picture are sound waves generated when the compressed gases expand from the muzzle. The bullet is centered in the air that it pushes out of the barrel by its piston action. Expanding turbulent gas behind the bullet gives it its acceleration.

If the study being made with this equipment involves transient conditions, it becomes necessary to synchronize the tripping of the spark gap with the particular point in the transient condition which it is desired to photograph. If the rate of change of the transient condition is very rapid, then this synchronization becomes a problem in electronic controls. Properly worked out, however, the equipment may be made very reliable, and can be made to show conditions at almost any previously chosen time in a transient condition lasting only a thousandth of a second or even less. The picture in Fig. 6 illustrates the method of allowing a sound wave to fall upon a microphone, thereby producing the impulse which triggers the flash.

Figure 7 is a shadowgraph of a bullet being discharged from a gun. Ordinary process films and contrast developers were used in making the shadowgraphs. Even if the film is placed at a distance of 20 feet from the spark gap, very good photographic exposures can be obtained, although the duration of the exposure time, as in this case, is only a fraction of a millionth of a second.

In many types of studies, shadowgraphs will make available practically all of the required information. If this is so, then the simplicity of the technique makes it highly preferable to the use of schlieren equipment. However, if the last possible detail of information regarding conditions and gradients is needed, then the shadowgraph is greatly lacking, and the use of schlieren equipment becomes almost mandatory.

The following bibliography covering the schlieren and shadowgraph methods is probably not entirely complete, some references being necessarily omitted because of military restrictions. However, it is hoped that this bibliography will afford a means of tying together the many loose ends of this rapidly expanding field. The authors wish to take this opportunity to thank Henriette Davidge for her excellent work in preparing this bibliography.

BIBLIOGRAPHY

Schlieren and Shadowgraph Method of Airflow Analysis

(1) J. Ackert, *Handbuch der Physik* (Verlagsbuchhandlung, Julius Springer, Berlin, 1927), Vol. VII, Pp. 308–342.

(2) H. S. Allen, Proc. Roy. Phil. Soc., Glasgow **33**, 71 (1902).

(3) *Annual Report of the National Physical Laboratory for 1933* (H. M. Stationery Office, London).

(4) F. Auerback and W. Hort, *Handbuch d. physik. u. techn. Mechanike* (Johann Ambrosius Barth, Leipzig, 1928), Vol. 5, p. 486.

(5) L. Bairstow, *Applied Aerodynamics* (Longmans Green, London, 1939).

(6) H. Bénard, Comptes rendus **147**, 970 (1908).

(7) G. B. Brown, Phil. Mag. **13**, 161 (1932).

(8) G. Burniston Brown, Proc. Phys. Soc. **47**, 703 (1935).

(9) W. Bucerius, Betriebsführung **11**, 33 (1932).

(10) W. Bucerius, Schornsteinfeger und Technik **57**, 65 (1932).

(11) A. Busemann, Luftwissen **9**, 173 (1942).

(12) A. Busemann, Z.V.D.I. **84**, 857 (1940).

(13) A. Busemann, Verh. d. III. Intern. Kongr. f. techn. Mech., Stockholm **I**, 282–86 (1930).

(14) C. Cranz, *Anwendung der elektrischen Momentphotographie auf die Untersuchung von Schusswaffen.* (Knapp, Halle, 1901).

(15) C. Cranz, *Lehrbuch der Ballistik* (Verlagsbuchhandlung, J. Springer; Berlin, 1927), second edition, Vol. 3, pp. 259, 264.

(16) C. Cranz and E. Bames, Zeits. f. angew. Chemie **36**, 76 (1923).

(17) C. Cranz and B. Glatzel, Ann. d. Physik **43**, 1186 (1914).

(18) C. Cranz, P. A. Günther, and F. Külp, Zeits. f. d. ges. Schiess- u. Sprengstoffwesen **9**, 61 (1914).

(19) C. Cranz, and P. A. Günther, Zeits. f. d. ges. Schiess- u. Sprengstoffwesen **7**, 317 (1912).

(20) C. Cranz and R. Kock, Ann. d. Physik **3**, 247 (1900).

(21) C. Cranz and F. Külp, Artillerist. M. **88**, 251 (1914).

(22) C. Cranz and H. Schardin, Zeits. f. Physik **56**, 147 (1929).

(23) S. Czapski, Zeits. f. Instrumentenk. **6**, 139 (1886).

(24) V. Dvorak, Ann. d. Physik **9**, 502 (1880).

(25) V. Dvorak, Zeits. f. physik. chem. Unterricht **21**, 17 (1908).

(26) E. Eckert and W. Weise, V.D.I. Forschung (Nov.-Dec. 1942).

(27) R. Emden, Ann. d. Physik **69**, 264 (1899).

(28) R. Emden, Wiedemanns Ann. **69**, 264, 426 (1899).

(29) S. Erk, Z.V.D.I. **77**, 1119 (1933).

(30) S. Exner, Repert. d. Physik **21**, 555 (1886).

(31) A. Ferri, Tech. Memo. No. 901, National Advisory Comm. for Aeronautics (July 1939).

(32) Léon Foucault, Ann. de l'observatoire imp. de Paris **5**, 197 (1859).

(33) Léon Foucault, *Recueil des travaux scientifiques* (Paris, 1878).

(34) Garside, Hall, and Townend, Nature **152**, 748 (1943).

(35) D. B. Gawthrop, Rev. Sci. Inst. **2**, 522 (1931).

(36) D. B. Gawthrop, J. Frank. Inst. **214**, 647 (1932).

(37) D. B. Gawthrop, W. C. F. Shepherd, and G. St. Perrolt, J. Frank. Inst. **211**, 67 (1931).

(38) C. W. M. Gregor, Metals and Alloys **4**, 19 (1933).

(39) C. W. M. Gregor, World Power **19**, 75 (1933).
(40) R. Hermann, Luffahrtforschung **19**, 201 (1942).
(41) R. H. Humphrey and R. S. Jane, Trans. Faraday Soc. **2**, 420 (1926).
(42) K. Kieser, Phot. Ind. **28**, 729 (1930).
(43) K. Kieser, Zeits. f. angew. Chemie **43**, 587 (1930).
(44) H. Klug, Ann. d. Physik **11**, 53 (1931).
(45) F. Kohlrausch, *Lehrbuch der praktischen Physik* (B. G. Teubner; Leipzig and Berlin, 1930), sixth edition, p. 304.
(46) A. Lafay, Comptes rendus **CL11**, 694 (1911).
(47) A. Lafay, Tech. Aéron **3**, 169 (1911).
(48) A. Lafay, Tech. Aéron **4**, 91 (1911).
(49) B. Lewis and G. VonElbe, *Combustion Flames and Explosions of Gases* (The University Press, Cambridge, England, 1938), pp. 149–155.
(50) W. Lindner, V.D.I. Forschungsheft No. 326 (V.D.I. Verlag, Berlin, 1930).
(51) F. Luft, Photo-Woche **2**, 24 (1933).
(52) E. Mach, Ann. d. Physik u. Chemie **41**, 140 (1890).
(53) E. Mach, Pogg. Ann. **159**, 330 (1876).
(54) E. Mach, Wiener Berichte **95**, 764 (1887).
(55) E. Mach, Wiener Berichte **98**, 41, 1303, 1310, 1327, and 1333 (1889).
(56) E. Mach and G. Gruss, Wiener Berichte **78**, 467 (1878).
(57) H. Mach, Forschung auf dem Gebiete des Ingen. **14**, 77 (1943).
(58) L. Mach, Zeits. f. Luftschiffahrt u. Physik der Atmosphäre **15**, 129 (1896).
(59) L. Mach, Wiener Berichte **105**, 605 (1896).
(60) L. Mach, Wiener Berichte **106**, 1025 (1897).
(61) H. Malz and Conrady, Autog. Metallbearb. **25**, 117 (1932).
(62) Hermann Malz, *Die Grenzen der Schneidegeschwindigkeit beim Brennschneiden. Diss.* (Berlin, 1932).
(63) A. Nadai, Proc. Am. Soc. Test. Mat. **31**, 2 (1931).
(64) *National Physical Laboratory Report for the Year 1935*, p. 179–180.
(65) R. A. Nelson, Phys. Rev. **23**, 94 (1924).
(66) A. V. Obermayer, Zeits. f. Luftschiffahrt u. Physik der Atmosphäre **15**, 120 (1896).
(67) A. V. Obermayer, Mitt. Gegenstände des Artillerie und Genie-Wesens 815 (1897).
(68) W. Oschatz, V.D.I. Forschung **11**, 296 (1940).
(69) W. Payman and H. Robenson, Safety in Mines Research Board Paper No. 18 and 29 (1926).
(70) W. Payman and Titman, Proc. Roy. Soc. **A152**, 418 (1935).
(71) W. Payman and D. B. Woodhead, Proc. Roy. Soc. Inst. **2**, 522 (1931).
(72) J. St. L. Philpot, Nature **141**, 283 (1938).
(73) S. Ch. Pramanik, Proc. Indian. Ass. Cult. Sci. **7**, 115 (1922).
(74) L. Prandtl, Physik. Zeits. **8**, 23 (1907).
(75) Proc. Roy. Soc. **A145** (1934).
(76) B. Ray, Proc. Ind. Assoc. Cult. Sci. **6**, 95 (1920).
(77) Reports and Memos of British Aero Research Committee No. 1349.
(78) Reports and Memos of British Aero Research Committee No. 1454.
(79) Reports and Memos of British Aero Research Committee No. 1767.
(80) Reports and Memos of British Aero Research Committee No. 1803.
(81) Rieckeherr, Kriegstechn. Zeits. **3**, 383, 439, 513 (1900).
(82) A. M. Rothrock and R. C. Spencer, National Advisory Committee of Aeronautics 24th Annual Report 213 (1938).
(83) A. M. Rothrock and R. C. Spencer, National Advisory Committee of Aeronautics 25th Annual Report 313 (1939).
(84) H. Schardin, V.D.I. Forschungsheft No. 367 (V.D.I. Verlag, Berlin, 1934).
(85) H. Schardin, Heizung und Lüftung 167 (1932).
(86) E. Schmidt, Forsch. Gebiete des Ingen **3**, 181 (1932).
(87) E. Schmidt, Z.V.D.I. **74**, 1487 (1930).
(88) P. Schrott, Kinotechn. **12**, 40 (1930).
(89) W. C. F. Shepherd, Bur. Mines Bull. No. 345 (1932).
(90) A. Stodola, *Dampf-und Gas-Turbinen* (Verlagsbuchhandlung, J. Springer, Berlin, 1922), 98–112.
(91) R. Straubel, in Winkelmann *Handbuch der Physik* (Johann Ambrosius Barth, Leipzig, 1906), Vol. VI, p. 485.
(92) T. Suhara *et al.*, Report of the Aero Research Institute Tokyo Imperial University **50**, 187 (1930).
(93) Harry Svensson, Kolloid Zeits. **87**, 181 (1939). Harry Svensson, Kolloid Zeits. **90**, 141 (1940).
(94) A. Tanakadate, Comptes rendus **151**, 211 (1910).
(95) E. P. Tawil, Comptes rendus **191**, 92, 998 (1930).
(96) H. G. Taylor and J. M. Waldram, J. Sci. Inst. **10**, 378 (1933).
(97) H. G. Taylor and J. M. Waldram, J. Sci. Inst. **11**, 31 (1934).
(98) Technical Memo 1066 N.A.C.A. (June, 1944).
(99) Terazawa, Yamazaki and Akishino, Tokyo Report of Aeronautical Research Institute **1**, 213 (1924).
(100) O. Tietjens, in W. Wien and F. Harms *Handbuch der Experimentalphysik* (Leipzig, 1931), Vol. I, p. 695.
(101) Arne Tiselius, Trans. Faraday Soc. **33**, 524 (1937).
(102) A. Toepler, Ann. d. Physik u. Chemie **131**, 33–35 (1867).
(103) A. Toepler, Arch. d. Artillerie- und Ingenieurwesens 485 (1887).
(104) A. Toepler, *Beobachtungen nach einer neuen optischen Methode* (Bonn, 1864).
(105) A. Toepler, "Beobachtungen nach einer neuen optischen Methode," *Ostwalds Klassiker der exakten Wissenschaften* No. 157 (Leipzig, 1906).
(106) A. Toepler, "Beobachtungen nach der Schlierenmethode," *Ostwalds Klassiker der exakten Wissenschaften* No. 158 (Leipzig, 1906).
(107) A. Toepler, Pogg. Ann. **127**, 556 (1866).
(108) A. Toepler, Pogg. Ann. **128**, 126 (1866).
(109) A. Toepler, Pogg. Ann. **131**, 33 and 180 (1867).
(110) A. Toepler, Pogg. Ann. **134**, 194 (1868).

(111) M. Toepler, Ann. d. Physik **27**, 1043 (1908).
(112) Townend, Aeronautical Research Committee, Reports and Memoranda No. 1349.
(113) H. C. H. Townend, J. Ae. Sci. **3**, 343 (1936).
(114) H. C. H. Townend, J. Sci. Inst. **11**, 184 (1934).
(115) H. C. H. Townend, Phil. Mag. **14**, 700 (1932).
(116) Townend, Reports and Memos of British Aero. Research Committee No. 1614 (1933).
(117) Townend, Reports and Memos of British Aero. Research Committee No. 1634 (1934).
(118) Vielle, Mém. des Poudres et Salpêtres **10**, 177 (1899).
(119) R. Wachsmuth, Ann. d. Physik **14**, 469 (1904).
(120) Weinhold, Zeits. f. physik. chemie Unterr. **21**, 281 (1908).
(121) O. Wiener, Wiedemanns Ann. **49**, 105 (1893).
(122) Melchior Wierz, *Beiträge zur Theorie der Lichtbahnen und Wellenflächen in heterogenen isotropen Medien. Diss.* (Rostock, 1901).
(123) R. W. Wood, Nature 62 (August 9, 1900).
(124) R. W. Wood, Phil. Mag. **48**, 218 (1899).
(125) R. W. Wood, Phil. Mag. **50**, 148 (1900).
(126) R. W. Wood, Phil. Mag. **52**, 589 (1901).
(127) R. W. Wood, *Physical Optics* (The Macmillan Company, London, 1905), p. 94.

A Physical Photometer for Afocal Optical Systems*

D. B. McRae
Hawk-Eye Works, Eastman Kodak Company, Rochester, New York
(Received May 21, 1945)

IN a previous paper[1] the author mentioned a photometer which had been built for the purpose of measuring the transmission of heavy and bulky optical instruments which could not be conveniently transported to a darkened room nor could they be set in place on ordinary photometer benches. Inasmuch as this instrument has proved valuable not only for measurements on large range finders, but for almost all types of afocal optical devices, a description of it seems worth while.

A few of the instruments to be measured were relatively simple telescopes in which the optical axis was a simple straight line, but in the majority of cases, the entrance and exit beams made very abrupt angles with each other and with the principal exterior surfaces of the instrument, while the entrance and exit pupils were anywhere from one to five feet apart. There was such a variety of sizes and shapes that it did not appear feasible to build a fixed or bench type of photometer to cope with them.

For this reason it was decided to build an electrical measuring instrument consisting of three mechanically independent parts:

I.—A light source which could be aligned with the entrance pupil of the device whose transmission was to be measured.

II.—A receiving device which could be aligned with the exit pupil of the device, using a photo-tube to determine the light passing through the device.

III.—A suitable meter or indicating device to read the output of the photo-tube mentioned above.

It was desired that the instrument be insensitive to daylight and ordinary artificial lighting in order to facilitate its use anywhere in the plant. This was accomplished by introducing a "chopper" or interrupter to modulate the beam from the light source, and the use of an a.c. amplifier which would greatly amplify voltages of "chopper" frequency, but would discriminate against those of 120 cycles or less.

The instrument as built makes use of the method of focused images described by Moffitt and Taylor[2] and discussed by the author in a previous paper.[1] Figure 1 shows schematically the arrangement of equipment. The projector, I, sends out a beam of collimated light originating from an intermittently illuminated spot on a diffuser, G_1. Receiver, II, uses a lens to form an image of the source in its focal plane, J, and the brightness of the image is determined by measuring the flux passing through a small fixed aperture, E. If a telescope is introduced between I and II, the image on J will be changed in size if the telescope has other than unit power, but the image brightness would be unchanged if the telescope were perfect. Actually the brightness is reduced by a factor which is the transmission factor of the telescope. The brightness of image is again determined by measuring the flux passing through the standard aperture, E, and the ratio of this reading to the reading without the telescope is the desired transmission factor.

THE LIGHT SOURCE OR PROJECTOR

As a light source, S, a 200-watt projector bulb is used. Condenser lens, L_1, concentrates light upon diffusing glass, G_1, which is masked to a small circular hole, P, in a metal plate. Immediately behind G_1 is placed a chopper, C. This consists of a large flat metal disk rotating about the axis X–X and having its plane perpendicular to the optical axis. The outer part of the disk has 18 equally spaced holes, each of diameter a little greater than the diameter of the condenser beam at the plane of C. The space between holes is equal to the diameter of the holes. The disk is rotated at a speed of approximately 1800 r.p.m. by a small induction motor and serves to pass and obstruct the light beam alternately 540 times per second. Lens, L_2, is a good quality projection lens which collimates the light from P. A commercial 2″×2″ lantern slide projector

* Paper presented at the 29th annual meeting of the Optical Society of America, New York, New York, October 20 and 21, 1944.

[1] D. B. McRae, J. Opt. Soc. Am. **33**, 229–243 (1943).

[2] G. W. Moffitt and P. B. Taylor, J. Opt. Soc. Am. **8**, 511–518 (1924).

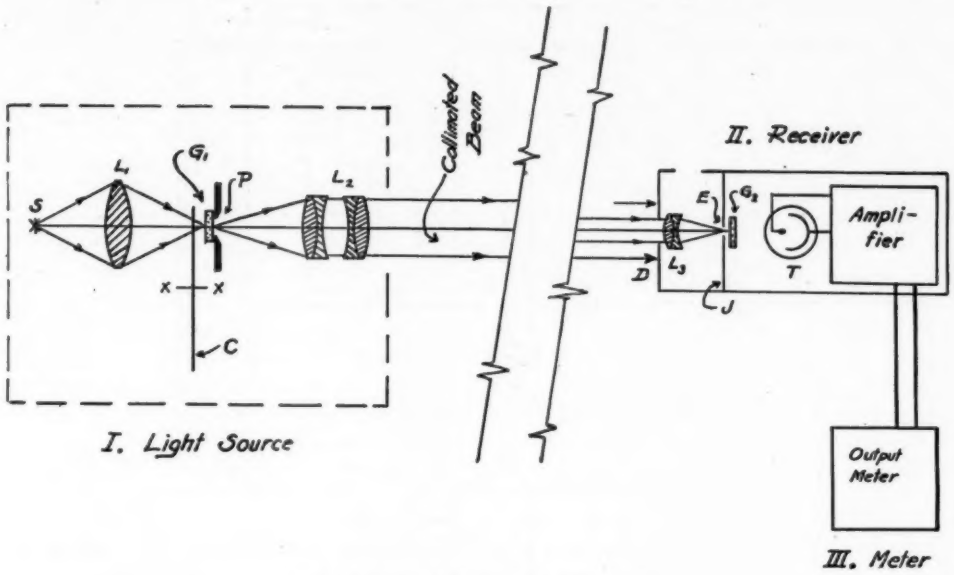


FIG. 1. Schematic layout of the photometer.

served as a basis for this part of the instrument. Part of the projector was moved forward enough to introduce the chopper, and the base was enlarged to permit the mounting of the chopper and motor.

THE RECEIVER

The receiver makes use of an arrangement which might be termed an "artificial eye" to view the image of P. A corrected lens, L_3, of about 1-inch focal length, is mounted so that a thin metal plate, J, lies in its principal focal plane. An iris diaphragm, D, is placed immediately in front of the lens to limit the diameter of the light beam entering L_3. A hole, E, of 1-mm diameter is drilled in the center of J, so that a portion of the light falling upon that area passes through E and is scattered over photo-cell, T, by diffusing glass, G_2. The excitation of the photo-cell is, therefore, proportional to the illumination falling on J. The output of the photo-cell is fed to an amplifier contained in the same metal box.

THE PHOTO-CELL AND AMPLIFIER

The circuit of the photo-cell, amplifier, and output meter is shown in Fig. 2. The output from a 926 photo-cell is amplified by a two-stage, resistance-capacity coupled amplifier consisting of a 6J7 vacuum tube followed by a 6SC7. Only one of the triodes in the envelope of the latter tube is used.

The output of the 6SC7 is carried to a rectifier and vacuum tube voltmeter (located in another metal box) by means of a shielded lead from condenser C-5 to the rectifier-input attenuator, R-9. The rectifier and voltmeter units utilize conventional circuits. The rectifier consists of a 6H6, with both sets of plates and cathodes connected together, and a resistance-capacity network to smooth the output. The voltmeter is a balanced bridge circuit consisting of two 6G6G tubes operated as triodes, and the equal resistors R-13 and R-14. The d.c. output from the rectifier is applied to the grid of one tube while the other grid is grounded. A 0–1 milliammeter is used at M to read the unbalance current, which is proportional to the voltage applied to the grid of the 6G6G. With no receiver output the meter is adjusted to zero reading by means of R-12. With any given receiver output the meter may be brought to full scale reading by means of R-9. These adjustments proved to be nearly independent although a large change of R-9

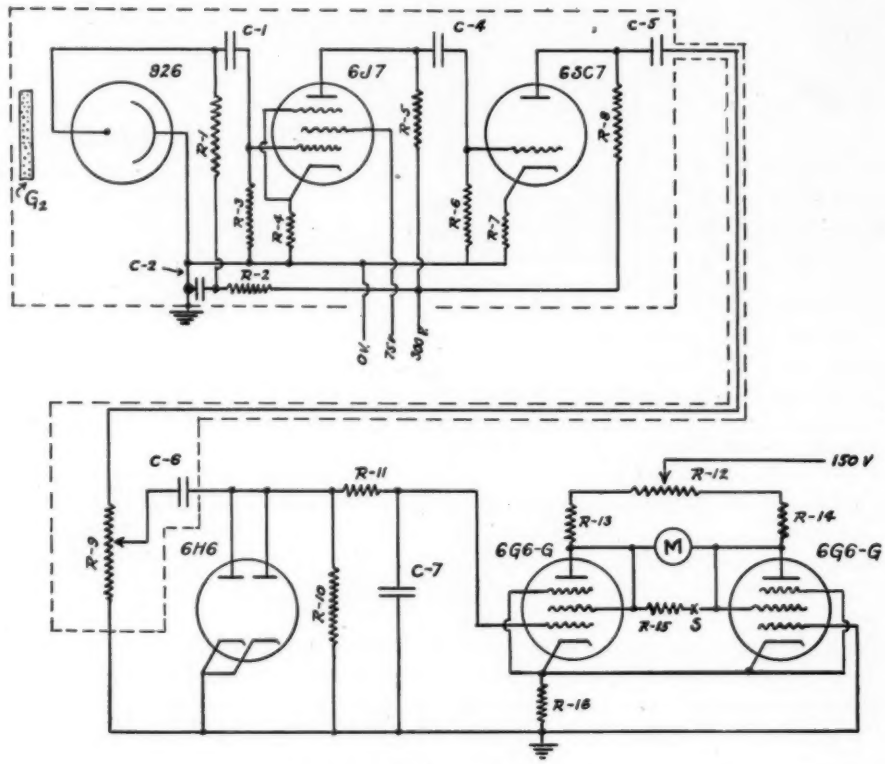


FIG. 2. Amplifier and meter circuits.

Resistor—Condenser Data for Fig. 2

R-1	10	megohms	*R*-9	1	megohm	*C*-1	0.001	microfarad
R-2	1	megohm	*R*-10	0.25	megohm	*C*-2	1	microfarad
R-3	1	megohm	*R*-11	0.5	megohm	*C*-3	25	microfarads formerly placed across *R*-4, but now omitted.
R-4	3000	ohms	*R*-12	550	ohms			
R-5	0.25	megohm	*R*-13	2500	ohms			
R-6	1	megohm	*R*-14	2500	ohms			
R-7	2000	ohms	*R*-15	25	ohms	*C*-4	0.001	microfarad
R-8	0.1	megohm	*R*-16	200	ohms	*C*-5	0.005	microfarad
						C-6	0.5	microfarad
						C-7	0.1	microfarad

requires a slight resetting of *R*-12. A shunt, *R*-15, with cut-out switch, *S*, proved useful to prevent extreme meter deflections during the "warm-up" period. The power supply is not shown in the diagram but is of conventional design with the three d.c. voltages required, namely 75, 150, and 300 volts, taken from taps on a voltage divider.

The rectifier, voltmeter, and power supply units are housed in a single metal box with the milliammeter, controls *R*-9 and *R*-12, meter shunting switch *S*, and the on-off switch for the power supply located on the front panel. Connections between this box and the receiver box are made by means of 12 feet of a special six-conductor cable. One wire of this cable has an individual shield and is used for the connection from *C*-5 to *R*-9. The remaining wires are enclosed in a common shield and are used to carry the ground, the 75- and 300-volt d.c. plate- and screen-supply voltages, and 6.3-volts a.c. for the heaters of the tubes. The length of this cable permits the receiver to be placed over the eyepiece of any range finder we have yet encountered, while the much heavier box containing the meter and power supply is set up on a convenient bench or table.

Interstage coupling in the amplifier by means

of condensers prevents the passage of any d.c. component from the photo-tube output. Stray daylight thus produces no amplifier response. It was originally planned to use an amplifier tuned to pass only the chopper frequency, but it would have been necessary to drive the chopper with a synchronous motor to maintain constant frequency, and such a motor was not readily available. It was found, however, that with the values of coupling condensers and grid resistors chosen, the response to 120 cycles from fluorescent lights in the room was low enough so that no visible meter movement resulted unless the actual image of such a light source were formed on the hole, *E* (Fig. 1), while signals of 540-cycle frequency were but slightly attenuated. If better discrimination were required, condensers *C*-1 and *C*-4 could be made smaller at the expense of reduced amplification.

Considerable care is necessary in the design of the electronic circuits for this type of instrument. It was found possible to secure good linearity of response in every unit except the rectifier, and even this has a characteristic which is quite straight except for a short curved portion near the beginning. As a result the over-all calibration of the detector is essentially the characteristic curve of the rectifier. As the intensity of the light falling on the photo-tube varies from one measurement to another, the signal level in the amplifier varies, and correspondingly, the setting of *R*-9 to give full-scale meter deflection varies. The calibration, however, remains the same for all settings of *R*-9, except for a short range near the grounded end. Settings in this region mean that the amplifier signal level is very high and the amplifier tubes are overloaded. The extreme lower portion of *R*-9 is, therefore, not used, the signal level being reduced by closing diaphragm, *D* (Fig. 1), until *R*-9 can be set somewhere in the upper region.

To determine the best operating conditions for the amplifier tubes, each was set up in a circuit similar to the one in which it was expected to operate. Voltmeters were connected to plate and grid. A battery with voltage divider was arranged so the grid could be set at various potentials above and below ground potential. Grid and plate voltages (actual voltage at the plate—not the plate supply voltage) were read for each setting and plotted. The values of the cathode resistor and the plate-load resistor could then be varied until zero signal voltage on the grid corresponded to a point at the center of the straight portion of the curve. The slope of the curve indicated the actual voltage amplification. The best operating conditions for triodes were quite easily obtained in this way. With pentodes, however, it was sometimes found necessary to make considerable changes in screen voltage from the published data before satisfactory performance could be obtained.

Ordinary filtering of the power supply was adequate for everything except the anode supply of the photo-tube. At this point a slight 120-cycle ripple caused trouble because it was impressed directly on the grid of the 6J7 and resulted in a steady half-scale reading of the meter. The load resistor of the photo-tube had a value of 10 megohms, so the addition of a simple decoupling filter *R*-2, *C*-2, with *R*-2 equal to 1 megohm and *C*-2 equal to 1 microfarad, was quite feasible. Such a filter theoretically should reduce 120-cycle ripple to about 0.13 percent of its original value. Actually it was reduced to a point where no meter deflection was observable.

CALIBRATION

The calibration of this type of instrument presented a few problems. The most fundamental method making use of the inverse-square law

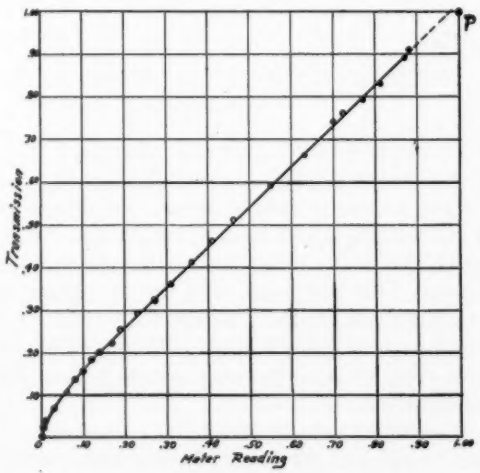


FIG. 3a. Calibration with neutral densities.

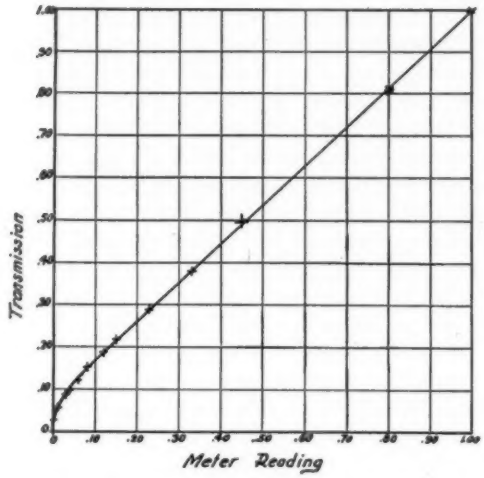


FIG. 3b. Calibration with measured apertures.

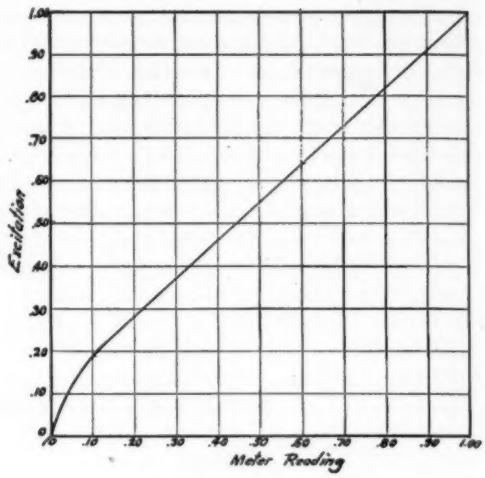


FIG. 4. Rectifier response.

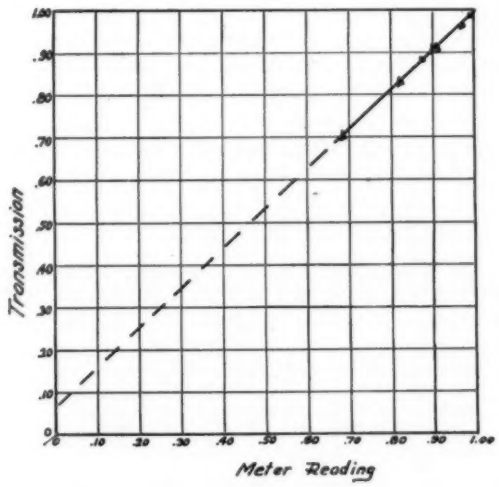


FIG. 3c. Calibration with tilted glass plate.

did not appear promising because of difficulties in maintaining a sufficiently high level of illumination and at the same time getting the chopper properly into the beam. The use of calibrated rotating-sector blades to reduce the intensity by known factors was not applicable because of the production of beats between sector blades and the chopper.

A set of 25 neutral density filters covering a density range of .04 to 1.41 (transmission range 0.91 to 0.04) was prepared and carefully measured for specular density on a high quality densitometer. Each filter was then placed in the beam and the meter reading noted, giving the calibration points shown by small circles in Fig. 3*a*. In each case the meter was adjusted to give full scale deflection just before placing the filter in the beam and checked for full scale deflection immediately after removing the filter. If the point marked P could be ignored, a plausible appearing straight line calibration would result for meter readings higher than 0.10. The curve between 0.10 and zero reading is explained as nonlinearity of the rectifier at low signal levels. The measured characteristic of the rectifier alone is shown in Fig. 4. The point P, however, cannot be ignored as it represents the initial reading of the meter with unobstructed beam. It is one of the two points whose position is known with certainty—the other being the zero reading with completely obstructed beam. The six upper points represent low density readings in which the error could be fairly large. An independent calibration was then carried out using a series of measured apertures. A beam of "chopped" light having uniform cross section was obtained by the arrangement shown in Fig. 5. Light from the projector is focused by lens L_4 onto a small spot on opal glass G_3. A diffusing light source of small area is thus formed at the center of G_3. To get more light at this point, diffuser G_1 was

removed from the projector. Lenses L_5 and L_6 are highly corrected telescope objectives of about 200-mm focal length and about 2-inch diameter. A parallel beam of light is formed between L_5 and L_6 and is brought to a focus on diffuser G_2 of the receiver. Lens L_3 and diaphragm, D, were removed during this test and aperture, E, was made large enough to be sure that it did not obstruct the converging beam. Uniformity of the beam between L_5 and L_6 was proved by exploring that region with a small aperture in a thin metal sheet and noting that the meter reading remained constant. A series of circular apertures of various sizes was prepared, the largest being a little smaller than the diameter of L_5 and L_6. The light passed by the largest was taken 100 percent, and the transmission of the others was considered to be the ratio of their areas to that of the largest. These known apertures gave the calibration points shown as crosses in Fig. 3*b*.

A third method of calibration was used particularly to obtain points between 0.90 and 1.00. A thin flat plate of good quality glass was set in the beam between projector and receiver, so that it could be tilted to allow light to pass through at various angles of incidence. The index of refraction of the glass was 1.5267. By use of Fresnels equations the reflection losses and the resulting transmissions were calculated for various angles of incidence and the results were corrected for the effect of multiple reflections between the two surfaces of the plate. Total transmission of light polarized both parallel and perpendicular to the plane of incidence was used because the instrument is independent of polarization.

In Table I the total transmission, $T\phi$, of this glass plate for various angles of incidence, ϕ, has been tabulated. Looking at the values of $T\phi$ in the second column we see that the maximum transmission is 0.917. It was possible to obtain calibration points for any desired number of points between this value and 1.000 by the following procedure. Instead of taking the receiver response as 100 percent with the glass plate removed from the beam, the transmission of the plate at normal incidence is taken as 100 percent. The plate is then turned to an angular position and the ratio of transmission at that angle to the transmission at normal incidence is taken. The value of these relative transmissions is given in the third column of the table. This method has proved to be a very convenient means of reducing the intensity of a light beam by small but accurately known amounts for this and other purposes. It has the great advantage that the angular settings of the plate are not very critical for either the $T\phi$ or the $T\phi/To$ values, very large changes in the angle giving rather small changes in transmission. It is necessary, however, that the plate be clean, and that it shall have been polished recently enough so that one may be sure that there is no surface tarnish to prevent accurate use of Fresnel's equations for reflection losses. The calibration points obtained by this method are shown as small triangles in Fig. 3*c*. The calibration points from the three methods have been combined in Fig. 6, and a straight line drawn to pass through P and as many of the calibration points as possible, giving greatest weight to the points determined by means of the tilted glass plate and the measured apertures. At the low end, the line is curved to follow the calibration points which were found to lie almost exactly on the rectifier characteristic. It is believed that the accuracy of measurements made with this instrument is better than ± 2 percent, provided that reasonable care is taken in the set-up.

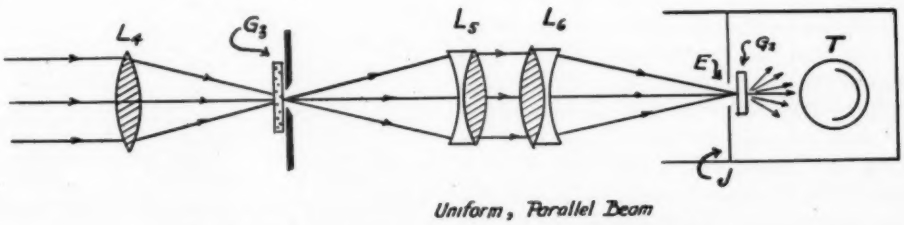


FIG. 5. Arrangement for calibrating with measured apertures.

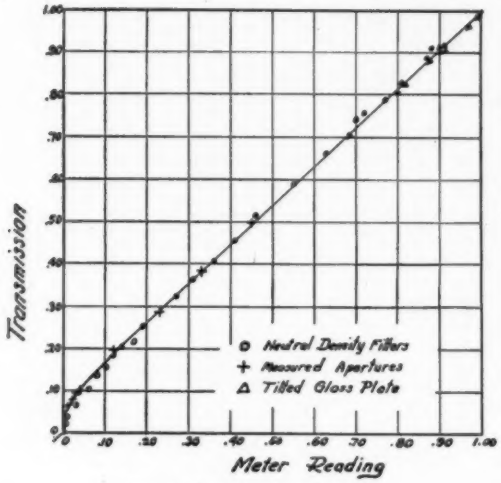


FIG. 6. Combined calibration.

USE AND APPLICATIONS

The use of the instrument to measure the transmission of a straight telescope is fairly simple. The projector is set up so that it sends the light beam through the telescope objective as nearly parallel to the optical axis as possible. This can usually be determined by inspecting the exit beam on a piece of ground glass, moving the latter back and forth and observing the symmetry of the patterns produced. The receiver is then placed in position at the eyepiece and moved until the exit pupil, or Ramsden disk, is focused on diaphragm D, after which the latter is closed down until the opening is definitely smaller than the exit pupil. This is indicated by the appearance of an annular ring of light concentric with the edge of the diaphragm. The telescope is then removed and without changing D, the light beam is permitted to enter lens L_3 and form an image of the effective light source, G_1, on the plane, J. By suitable orientation of the receiver this image is centered on the small hole, E. It is necessary that the image be larger than E, so that the flux through E may be a measure of the image brightness. The meter is next adjusted for full scale deflection by use of gain control R-9, and the zero checked by obstructing the beam and readjusting R-12 if necessary. The telescope is put back in position, and the meter read. The transmission is then determined from the calibration curve.

The procedure is similar in the case of more complicated telescopes and range finders with bent axes, except that in these cases the projector and receiver must be supported in such manner that the beam can be sent in along the axis and properly picked up by the receiver. With some range finders the set-up has required considerable ingenuity. In such cases the comparison reading cannot be made by the simple method of lifting the telescope out of position but must be made by moving both projector and receiver to a position where they can be pointed towards each other. An auxiliary plane, front-surfaced mirror mounted to swivel around the projector lens L_2, has proved useful in getting the light beam into some instruments. The mirror must, of course, also be used in the comparison measurement in this case.

Flat plates of glass and neutral density filters are the simplest of all things to measure. They are simply moved in and out of the beam. The transmission of non-dispersive deflecting prisms and the reflectivity of mirrors are also quite easy to measure. The receiver is aimed at the projector and set for full scale meter deflection and then swung to the required angle to pick up the beam as it comes from the prism or mirror.

This instrument was designed for measurements on afocal devices, but the transmissions of photographic lenses have been measured on occasion by the device of putting two lenses together to make a telescope. Only the product of the two transmissions can be measured, but if three lenses are available, they can be paired in three different ways. In this way three equations are obtained and the individual transmissions may be calculated. A photometer of the Sabine[3] type is preferable for this work, however.

TABLE I.

ϕ	$T\phi$	$T\phi/To$
0°	0.917	1.000
10°	0.917	1.000
20°	0.916	1.000
30°	0.914	0.997
40°	0.906	0.988
50°	0.884	0.964
60°	0.829	0.904

[3] D. B. McRae, J. Opt. Soc. Am. **33**, 234 (1943).

COMMENTS

This instrument was designed for the purpose of making occasional measurements on a variety of afocal, non-dispersive, optical devices. Versatility was considered of greater importance than speed of use on any particular job, and it was felt that because of the uncertainties involved in lining up projector, telescope, and receiver an accuracy of 2 percent was quite adequate for our purposes, and that greater accuracy was not worth the added expense.

The choice of parts, meter, amplifier tubes, etc., was largely the result of attempting to use what was on hand rather than wait for many months to obtain something which might ordinarily seem more desirable.

Some rather definite ideas have arisen during the use of this instrument which would be desirable to incorporate in a second model if one were built. The amplifier circuit should be balanced, rather than single sided. The extra circuit could very well use a photo-tube which would be excited by a portion of the light from the projector. This second photo-tube and amplifier circuit need not be in the receiver box but could be incorporated in the projector and be connected to the proper place in the circuit by a cable. Trouble from variations in voltage applied to lamp, S, would be reduced and fewer comparison readings would be required.

It would also be desirable to incorporate a calibrated attenuator in the amplifier output circuit in addition to the control R-9, so that the output might be changed by a factor of precisely five or ten. Linear rectifier circuits are now said to be in existence and if these are reliable enough, it might be possible to get a straight line characteristic and read transmissions direct from the meter. This could, of course, be accomplished by putting a special scale on the meter.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 8 AUGUST, 1945

The Effective Aperture of a Photographic Objective

R. KINGSLAKE
Eastman Kodak Company, Rochester, New York
(Received May 26, 1945)

IT is a well-known fact that no matter what the object distance may be, the illumination E at the center of the field of a real optical image is given by

$$E=\pi kB \sin^2 U', \qquad (1)$$

where k is the transmission factor of the lens, B is the brightness of the object being photographed, and U' is the slope angle of the extreme marginal ray proceeding from the lens to the image.

The "speed" of a lens is, therefore, proportional to the product $(k \sin^2 U')$, and there is much to be said for the introduction of some system of speed-markings on lenses, based on this principle. The closest approach is found in the "numerical aperture" (N.A.) engraved on a microscope objective, which is equal to $\sin U'$, but even then no account is taken of the transmission factor k. For a great many years, the photographer has based all his exposure times on the "f-number" of the lens, by which is meant the ratio of the focal length, f, to the diameter of the entrance pupil, D. For a very distant object, the f-number is equal to $1/(2 \sin U')$, and for this special case, ignoring the transmission factor, the lens speed (or image illumination) is inversely proportional to the square of the f-number.

If the object is not at infinity, the lens speed is, of course, still measured by the sine of the emerging ray slope U'. It has therefore become customary to use the term "effective f-number," defined by $1/(2 \sin U')$, as a measure of the lens speed for a near object.

Unfortunately for a thick lens there is no simple relation whatever between the ordinary (or infinity) f-number, and the "effective f-number" at any specified object distance. A few examples will indicate why this is the case.

EXAMPLE 1. A THIN LENS USED AT A MAGNIFICATION m

For parallel light, the f-number is given by f/D. For a near object, the effective f-number is obviously equal to $f(1+m)/D$ (Fig. 1). Thus the effective f-number is larger than the infinity f-number by a factor $(1+m)$. The effective f-number of a thin $f/8$ lens used at equal conjugates is therefore $f/16$, and the speed of the lens will be only one-quarter of its speed for a distant object.

EXAMPLE 2. AN UNSYMMETRICAL LENS TURNED END-FOR-END

In an unsymmetrical lens, such as a telephoto, the diameters of the entrance and exit pupils are quite different, but the focal length is the same no matter which end of the system is receiving parallel light (Fig. 2). Hence the f-number for a very distant object may be two or three times as great when the lens is turned end-for-end. This may be regarded as an extreme example of a change in conjugate distances, the object being brought up from infinity to the anterior focal plane of the lens.

EXAMPLE 3. A 1:1 PRINTER SYSTEM WITH FIELD FLATTENERS

A 1:1 printer can be made up of two identical systems face to face, each with a negative field-flattener lens close to the film, and parallel light between the two halves (Fig. 3). Each half may

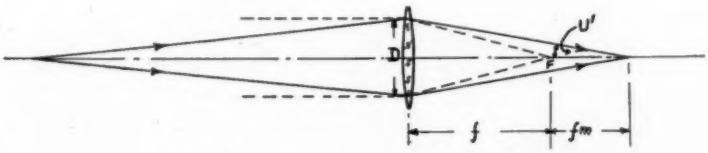


FIG. 1. A thin lens used at a magnification m.

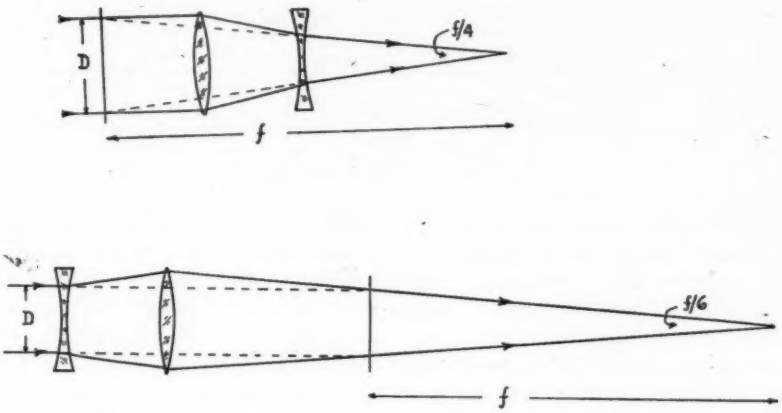


FIG. 2. The change in aperture on turning an unsymmetrical lens end-for-end.

have a focal length of, say, 4 inches and an aperture of $f/3$, the central diaphragm having a diameter of 4/3 inches. However, the focal length of the complete system may be perhaps 3 inches, and the entrance pupil diameter may be only about 3/4 inch, giving an over-all relative aperture for a distant object of $f/4$. Thus in this case the speed of the lens at unit magnification is somewhat greater than its speed with a distant object. It should be noticed that when such a lens is used with parallel light, the aperture stop is the first field lens, but this lens acts as a field stop when the system is used at 1:1.

EXAMPLE 4. AN AFOCAL 1:1 PRINTER

In some types of motion-picture printer, especially for use with lenticular color-film, the lens consists of two identical components mounted so that their adjacent focal points coincide (Fig. 4). Light from any point in the object plane is rendered parallel by the first component, and brought to a focus by the second component in its posterior focal plane.

If each component has a high relative aperture, such as $f/2$, the effective aperture of the printer, under service conditions, is of course $f/2$. However, as it is an afocal system, the over-all focal length of the whole system is infinite, and the f-number for a distant object is therefore infinite, and quite meaningless.

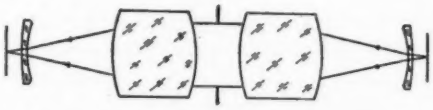


FIG. 3. A 1:1 printer with field flatteners.

EXAMPLE 5. A LONG LENS HAVING A SMALL REAR WINDOW

In some modern types of high aperture lens, the back focus is short and the rear window is much smaller than the front window. In such cases, it is often the rear window that determines the effective aperture of the lens (Fig. 5). To take a specific instance, suppose the back focus is half the focal length, and the infinity f-number is 2.0, the diameter of the rear window is therefore $f/4$. Then at a magnification of 1:4, the effective aperture will be $f/3$; and at unit magnification it will be $f/6$. The drop in speed as the object comes closer to the lens is thus much more rapid with a lens of this type than with a thin lens.

To show how meaningless is the infinity f-number in such a case, we notice that in Fig. 5, the front window may be drastically reduced in size without in any way affecting the lens speed with a near object, but this reduction would immediately raise the f-number of the lens defined by the simple ratio f/D.

These five examples should indicate that there is no direct relationship between the working speed of a thick lens at finite magnification and f-number as ordinarily defined. Lenses which are

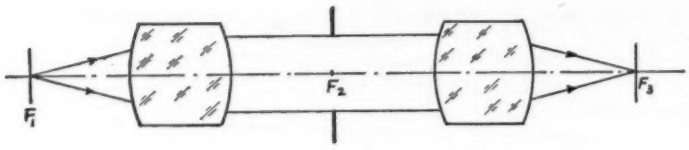


FIG. 4. An afocal 1:1 printer.

substantially symmetrical and rather short compared with their focal length can be regarded as "thin" lenses, and the rule given under Example 1 above may be used. This applies to all ordinary camera, process, and enlarging lenses, and some types of fixed-magnification projection printers. For a lens having a large front window and a small rear window, the speed of the lens drops faster than the thin-lens rule indicates as the object is brought closer to the lens. On the other hand, if the front window is smaller than the remaining parts of the lens, the speed will drop at a much slower rate, and it may even increase.

In all these special cases, especially where very long lenses are used at a fixed magnification, it is strongly urged that the lenses be engraved with their effective aperture at a specified magnification, and not with the ordinary *f*-number for a distant object. The iris diaphragm markings must, of course, be engraved in the same system.

A much more difficult situation arises if a long

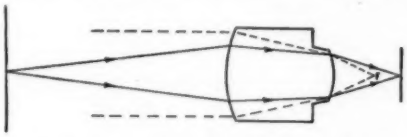


FIG. 5. A long lens with a small rear window.

lens is used over a range of magnifications. In this case, it may be desirable to supply a table of effective speeds against magnification, either engraved on the lens barrel or printed on a card fastened in the lid of the lens box.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 8 AUGUST, 1945

Uses of Retardation Plates in Spectrophotometry

III. Measurement of Dichroic Samples

GEORGE L. BUC AND E. I. STEARNS
Research Department, Calco Chemical Division, American Cyanamid Company, Bound Brook, New Jersey

(Received April 6, 1945)

A theory of retardation plate flickering is presented. The behavior of rotating plane of polarization spectrophotometers with dichroic samples is formulated. It is shown that polarized beam spectrophotometers modified by the addition of a thick retardation plate will duplicate the results of absorption measurements on dichroic samples obtained with unpolarized light.

A DICHROIC sample is one in which the absorption of polarized light is a function of the orientation. Transmission samples do not ordinarily exhibit dichroism although Polaroid and many dyed Cellophane samples are dichroic. Reflectance samples exhibit some degree of dichroism much more frequently than do transmission samples. Dichroism is especially prevalent in some weaves of synthetic fibers such as a satin weave of filament rayon (Fig. 1).

In general, polarizing spectrophotometers do not measure the absorption of a dichroic sample independently of the azimuth of the sample. The average of two measurements 90° removed in azimuth gives the absorption for unpolarized light.[1] The difficulties involved in obtaining a single measurement of absorption independent of polarization with methods heretofore available have recently been cited.[2]

The present report discusses the two types of polarization flickering commonly used in automatically recording spectrophotometers, presents the theory of measurement of dichroic samples with the two types and describes a method whereby a single measurement of absorption in polarized light will duplicate results obtained with unpolarized light.

The basic principle of polarizing spectrophotometers is to split an essentially monochromatic beam of light into two beams whose intensity ratio is governed by the angular position of a polarizing optic and then reduce the intensity of the stronger beam by the action of an absorbing sample (in transmission or reflectance) until it is equal to the weaker beam.[2] The equality of the two beams is judged by the eye[3] or by photoelectric means and the position of the polarizing optic gives a measure of the transmitting or reflecting powers of the sample. The photoelectric means of judging the equality of the beams may involve flickering the beams and in the more elaborate modifications any inequality of the beam gives rise to a rotation of the photometering optic tending to equalize the beams while a pen coupled with the photometering optic plots the absorption.[4–6]

Flickering has been accomplished mechanically[7] and by two polarization methods. One polarization method involves a rotating Rochon[4] and the underlying theory has been presented.[8] The other polarization method involves a rotating retardation plate followed by a fixed Rochon[5] for which the theory is presented below.

The retardation plate flickering spectrophotometer employs, in order, a Wollaston, a rotating retardation plate, and a fixed Rochon prism as the flickering optics. When the fixed Rochon is set to pass one of the Wollaston beams the intensities of the two beams after passing through the fixed Rochon are from Eq. (1), Part I,

$$I_1 = a_1^2[1-\sin^2 A \sin^2 (D/2)], \qquad (1)$$
$$I_2 = a_2^2[\sin^2 A \sin^2 (D/2)].$$

I_1 is the intensity of the reference beam, I_2 is

[1] R. D. Nutting, J. Opt. Soc. Am. **25**, 211 (1935).
[2] Committee on Colorimetry, J. Opt. Soc. Am. **34**, 183 (1944).
[3] K. S. Gibson, *Measurement of Radiant Energy* (McGraw-Hill Book Company, Inc., New York, 1937), pp. 326–52.
[4] O. W. Pineo, U.S. Pat. No. 2,107,836 (1938).
[5] O. W. Pineo, U.S. Pat. No. 2,126,410 (1938).
[6] W. R. Brode and C. H. Jones, J. Opt. Soc. Am. **31**, 743–49 (1941).
[7] A. C. Hardy, J. Opt. Soc. Am. **18**, 96 (1929).
[8] A. C. Hardy, J. Opt. Soc. Am. **25**, 305 (1935).

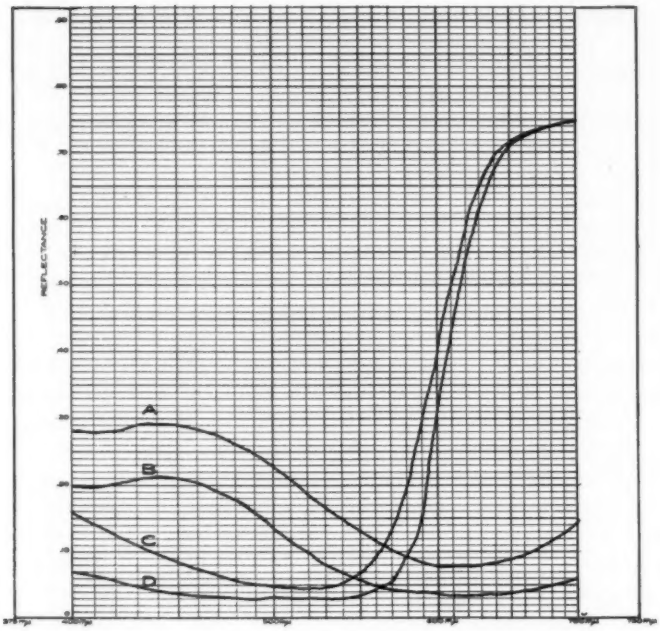


FIG. 1. Illustration of dichroism of reflectance samples. The samples are both satin weave filament viscose rayon and were measured at two azimuths 90° apart on a spectrophotometer with fixed planes of polarization. *A*, *B* are 1 percent dyeings—Benzo fast blue FR (CI. 533); *C*, *D* are ½ percent dyeings —Benzopurpurine 4B (CI. 448).

the intensity of the sample beam before striking the sample, and a_1^2 and a_2^2 are the intensities of the reference and sample beam leaving the Wollaston. A is the angle of the optical axis of the retardation plate with the plane of polarization of the Wollaston sample beam and D is the retardation angle of the plate. The ratio of a_1^2/a_2^2 is governed by the angle r that the photometering Rochon makes with the Wollaston in accordance with the usual $\tan^2 r$ relation. After attenuation by the sample the intensity of the sample beam may be written as

$$I_s = Ta_2^2[\sin^2 2A \sin^2 (D/2)]. \qquad (2)$$

The sample and reference beams combine in the integrating sphere to give a signal equal to their sum

$$S = I_1 + I_s = a_1^2 + (T a_2^2 - a_1^2) \sin^2 2A \sin^2 (D/2). \qquad (3)$$

If the plate is rotated the a.c. signal from the photo-cell will be

$$(T a_2^2 - a_1^2) \sin^2 2A \sin^2 (D/2). \qquad (4)$$

The phase of the a.c. signal and hence the direction of rotation of the photometering Rochon is controlled by the relative magnitudes of $T a_2^2$ and a_1^2. When $T a_2^2 = a_1^2$ the a.c. signal vanishes and the machine balances. For this case

$$T = a_1^2/a_2^2 = \tan^2 r. \qquad (5)$$

The amplitude of the a.c. signal has a factor $\sin^2 (D/2)$. This has a value of ½ for a quarter-wave plate, 1 for a half-wave plate, 0 for a full-wave plate. A retardation plate will produce flickering provided $\sin^2 (D/2) \neq 0$ but is most efficient for a half-wave plate. Since the value of $\sin^2 (D/2)$, averaged over a large angle, D, is ½, a sufficiently thick retardation plate will give 50 percent flickering for a finite slit.

The rotating Rochon and the rotating retardation plate types of polarization flickering result in different kinds of light beams striking the

sample. In Rochon flickering the plane of polarization of the beam rotates while in retardation plate flickering, the plane of polarization is stationary.

The operation of a rotating Rochon flickering spectrophotometer has been formulated for a non-polarizing sample.[8] The explanation of the unsatisfactory behavior of this type of flickering with a dichroic sample[9] requires a slight extension of the previously developed theory. The behavior of a retardation plate flickering spectrophotometer with dichroic samples is in accord with the published theory for stationary planes of polarization.[3]

Assuming a perfectly polarizing sample it is necessary to use for sample transmission the value $\cos^2(P-Q)$ in the formula previously derived.[8] This yields the following expression for the photo-cell current if photo-cell response is assumed linear with light intensity and independent of light polarization.

$$S=\frac{1+\sin^2 r}{2}-\frac{\cos^2 r}{4}\cos 2Q$$
$$+\cos 2P\left(\sin^2 r-\frac{\cos^2 r}{2}\right)$$
$$+\cos(2P-2Q)\frac{\cos^2 r}{2}$$
$$-\cos(4P-2Q)\frac{\cos^2 r}{4}. \quad (6)$$

S is the photo-cell current. The angles are all measured from the plane of polarization of the sample beam at the Wollaston prism; r is angle of the photometering Rochon, P is the angle of the flickering Rochon, and Q is angle of the polarizing sample. The first line of Eq. (6) is the d.c. photo-cell current and is not amplified. The second line is the normal balancing current. The third line represents a current of the same frequency as the balancing current but out of phase. The fourth line represents a double frequency variable phase current which can be tuned out. The third line term or out of phase signal is responsible for poor operation. It should be noted that for the cases of $Q=0°$ and $90°$ that the out of phase signal vanishes and the machine operates satisfactorily. Not only must two positions 90° apart be averaged to get the absorption for unpolarized light but the two positions must each be carefully chosen to get good machine operation.

The problem of getting the absorption of dichroic samples in a single measurement may be solved by the use of a thick retardation plate as is shown by the following considerations. In the fixed plane of polarization type of spectrophotometer suppose that a thick retardation plate be placed after the last polarizing optic and ahead of the sample with its optic axis at 45° to the plane of polarization of the beam. Let a Nicol be the sample in question. Then from Eq. (1), Part I, the transmission (T_c) of the Nicol and plate will be

$$T_c=\cos^2(45°-Q)-\sin 2Q\sin^2(D/2). \quad (7)$$

Further suppose that within one slit width of the spectrophotometer in question $D/2$ goes through many multiples of π. This condition can always be fulfilled for any finite slit width with a sufficiently thick piece of calcite as a retardation plate. Then the average value of $\sin^2(D/2)$ will be $\frac{1}{2}$ and substituting this value in Eq. (7) it is found that the transmission (T_c) becomes $\frac{1}{2}$ and independent of Q. One half is the value for the transmission of the Nicol for unpolarized light if reflectances at interfaces and losses caused by absorption be ignored. Thus a properly placed retardation plate results in a Nicol yielding a transmission value equivalent to that obtained with unpolarized light. Another way of considering it is that if $D/2$ goes through many radians within the range of wave-lengths comprised in one slit width, then the Nicol is viewed with light having such a mixture of plane, elliptical, and circular polarization that it is effectively viewed with unpolarized light.

The same optic, a thick retardation plate, may be attached at 45° to the rotating Rochon prism of the second type of spectrophotometer. Under these conditions the value of the transmission for a Nicol is not $\cos^2(P-Q)$ as in Eq. (6) but is $\frac{1}{2}$. The polarization effect which introduces the out of phase term in Eq. (6) is thereby eliminated and the spectrophotometer becomes operative at all sample azimuths and

[9] H. J. Keegan, Inter-Society Color Council Minutes, p. 21 (March 1, 1944).

gives the value which would be obtained with unpolarized light.

In adjusting the thick retardation plate a sufficiently critical criterion of correct angular adjustment is that the transmission of a Polaroid is independent of its azimuth. If a Polaroid can be measured satisfactorily in transmission then any reflectance sample will also be measured satisfactorily as few, if any, reflectance samples exhibit the same degree of dichroism as Polaroid. A calcite plate about 2 mm thick will work satisfactorily in the visual range for 10-mμ slits and for narrower slit widths a correspondingly thicker plate may be chosen.

Methods involving quarter-wave plates[10] and rotation of the sample[11] to eliminate the effects of sample dichroism have been proposed but are not as convenient or efficient as the method given here. Depolarization schemes give scattering and a consequent intolerable loss of light.

If it is desired to evaluate dichroism and the spectrophotometer has been adapted with a retardation plate as indicated above, a fixed Nicol placed after the retardation plate converts the adapted instrument to a fixed plane of polarization spectrophotometer. Measurements made with varying angles between the Nicol and sample will then show the dependence of absorption upon polarization orientation.

[10] O. W. Pineo, U.S. Pat. No. 2,189,270 (1940).

[11] E. M. Shelton and R. L. Emerson, Ind. Eng. Chem. Anal. Ed. **4**, 248 (1932).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 8 AUGUST, 1945

The Statistical Properties of Unpolarized Light

HENRY HURWITZ, JR.
Cornell University, Ithaca, New York

(Received November 1, 1944)

In a beam of monochromatic unpolarized light the electric field vector at a point traces out an ellipse whose size, eccentricity, and orientation are slowly varying functions of time. The statistical properties of the parameters of this ellipse are investigated. It is shown that the quantity S which is defined as twice the product of the principle axes of the ellipse divided by the sum of the squares is uniformly distributed between zero and one. It therefore has median value $\frac{1}{2}$ which corresponds to a ratio of minor to major axis equal to .268. Hence fairly thin ellipses predominate. The square root of the sum of the squares of the semi-major and semi-minor axes, R, is statistically independent of S and has the distribution function $(r^3/2p^4)\exp(-r^2/2p^2)$ where $2p^2$ is the average value of R^2.

INTRODUCTION

MANY textbooks and teachers of optics are rather vague in their descriptions of unpolarized light. They pass over the subject with a cryptic statement of the sort "unpolarized light is a mixture of plane-polarized light in all directions" or "unpolarized light is a random mixture of light with all kinds of polarizations." Students find difficulty in obtaining from these brief remarks a clear picture of how the electric field vector at a point varies as a function of time. In this note we shall perform some simple statistical calculations whose results might serve as a pedagogical aid in making the situation clear to students.

By natural, or unpolarized light we shall mean radiation which is emitted by molecules which have been randomly excited (e.g. thermally, or by an electrical discharge), and which are not in a region of space where large external fields provide a preferred direction. We shall assume that the beam of light under consideration is monochromatic in the sense that its line breadth is small compared to the central frequency ν_0. It is immaterial whether the monochromatic character arises because the light has been passed through a narrow filter, or because it consists of a single spectral line. In order to avoid the necessity of considering quantum effects we shall limit the discussion to fairly intense beams. This means effectively that the number of photons reaching the detector in a time interval equal to the reciprocal line breadth must be large compared to one.

As will be seen from the formulas given below, a beam of unpolarized monochromatic light would, if observed over a time interval short compared to the reciprocal of the line breadth, appear to be elliptically polarized. But the type of polarization changes continually so that in any experiment which requires a longer period of observation one measures only the average effect of a large variety of different polarizations. In other words, in a beam of unpolarized monochromatic light the electric vector at a point traces out an ellipse whose size, eccentricity, and orientation gradually change. After a time somewhat larger than the reciprocal of the line breadth the new ellipse is completely unrelated to the original one. The purpose of our calculation is to find the statistical properties of the various parameters of the "instantaneous ellipses" traced out by the electric vector.

STATISTICAL FOUNDATIONS

Our statistical calculations will be based on the following mathematical representation for a monochromatic beam of unpolarized light:[1]

$$\begin{aligned} E_x &= X(t)\cos\omega_0 t + U(t)\sin\omega_0 t \\ &= A(t)\cos(\omega_0 t - \delta_x(t)), \\ E_y &= Y(t)\cos\omega_0 t + V(t)\sin\omega_0 t \\ &= B(t)\cos(\omega_0 t - \delta_y(t)), \\ A(t) &= [X^2(t)+U^2(t)]^{\frac{1}{2}}; \\ B(t) &= [Y^2(t)+V^2(t)]^{\frac{1}{2}}; \\ \delta_x(t) &= \tan^{-1}[U(t)/X(t)]; \\ \delta_y(t) &= \tan^{-1}[V(t)/Y(t)]; \quad \omega_0 = 2\pi\nu_0. \end{aligned} \tag{1}$$

[1] We assume that the electric field is constant over the part of the cross section of the beam under consideration.

Here E_x and E_y are the x and y components of the electric field at a point P in the beam. The quantities $X(t)$, $Y(t)$, $U(t)$, and $V(t)$ are statistically independent Gaussianly distributed random functions. The probability that a particular one of them has a value between x and $x+dx$ is

$$P_X(x)dx=\frac{1}{p}P_G\left(\frac{x}{p}\right)dx$$

$$=\frac{1}{(2\pi)^{\frac{1}{2}}}\exp\left(-\frac{x^2}{2p^2}\right)\frac{dx}{p}, \quad (2)$$

where p^2 is the average of the square of each one of these quantities. In accordance with Eq. (2) and the independence of $X(t)$, $Y(t)$, $U(t)$, and $V(t)$, the quantities $A(t)$, $B(t)$, $\delta_x(t)$, and $\delta_y(t)$ are all independent. Furthermore $\delta_x(t)$ and $\delta_y(t)$ are uniformly distributed between 0 and 2π, and $A(t)$ and $B(t)$ are distributed according to the law

$$P_A(x)dx=P_B(x)dx=\frac{x}{p^2}\exp\left(-\frac{x^2}{2p^2}\right)dx. \quad (3)$$

(In this notation $P_A(x)dx$ is the probability that $A(t)$ lies between x and $x+dx$, etc.) The eight functions $X(t)$, $Y(t)$, $U(t)$, $V(t)$, $A(t)$, $B(t)$, $\delta_x(t)$, and $\delta_y(t)$ vary slowly compared to $\cos\omega_0 t$ and $\sin\omega_0 t$. Thus the values of one of these functions at two different times are not statistically independent unless the time difference is large compared to the reciprocal line breadth. For smaller time differences the correlation depends on the shape of the line.

We shall indicate two different methods of obtaining the representation given in Eqs. (1) to (3). The first is based on a simple classical model of light emission while the second makes use of a natural independence assumption.

In the classical model which we shall employ in the first derivation the molecules are assumed to emit radiation in the form of wave trains which contain only Fourier components with frequencies close to ν_0 and which last for a time interval of order of magnitude of the reciprocal line breadth. (In accordance with quantum theory the pulses may be imagined to contain an amount of energy equal to $h\nu_0$.) The nature of the polarization of the individual pulses is unimportant so long as it is sufficiently random. For example, if the pulses are plane polarized, then all planes of polarization must be equally probable, while if the pulses are circularly polarized, right- and left-handed polarizations must be equally likely. A typical field component in a pulse will be equal to $f(t-t_i)$ where t_i is the time when the pulse begins. Expanding $f(t)$ in a Fourier series in the long time interval 0 to T we may write

$$f(t)=\sum_{k=0}^{\infty}(\alpha_k\cos\omega_k t+\beta_k\sin\omega_k t), \quad (4)$$

$$\omega_k=2\pi k/T.$$

The quantity $\alpha_k^2+\beta_k^2$ is small if ω_k differs from ω_0 by an amount large compared to the line breadth. The problem of finding the resultant of a large number of pulses of the sort described above has been studied in great detail in other connections.[2] Therefore without going into details about the mathematical calculation we shall simply quote the result. From the fact that the times t_i when the individual pulses are emitted are randomly distributed it can be shown that the x and y components of the electric field at a point in the beam may be written in the following form:

$$E_x=\sum_{k=0}^{\infty}p_k(X_k\cos\omega_k t+U_k\sin\omega_k t),$$

$$E_y=\sum_{k=0}^{\infty}p_k(Y_k\cos\omega_k t+V_k\sin\omega_k t). \quad (5)$$

---

[2] H. Hurwitz and M. Kac, Ann. Math. Stat. **15**, 173 (1944); S. O. Rice, Bell Sys. Tech. J. **23**, 282 (1944), **24**, 46 (1945). Some insight as to why the quantities in Eq. (5) are Gaussianly distributed can be had from the following considerations. Let

$$F(t)=\sum_1^N f(t-t_i)=\sum_{k=0}^{N}(a_k\cos\omega_k t+b_k\sin\omega_k t).$$

From Eq. (4)

$$a_k=\sum_1^N(\alpha_k^2+\beta_k^2)^{\frac{1}{2}}\cos(\omega_k t_i-\varphi_k),$$

$$b_k=\sum_1^N(\alpha_k^2+\beta_k^2)^{\frac{1}{2}}\sin(\omega_k t_i-\varphi_k), \quad \varphi_k=\tan^{-1}\frac{\beta_k}{\alpha_k}.$$

If the t_i's are randomly distributed the quantities a_k and b_k can be regarded as the two components of a vector which is the sum of N two-dimensional vectors of constant length but random direction. Hence from the familiar results of the problem of "random walk" we can immediately conclude that a_k and b_k are Gaussianly distributed in the limit that N becomes infinite.

The quantities p_k^2 are proportional to $\alpha_k^2+\beta_k^2$ and are equal to $2\pi/c$ times the radiant-energy current in the beam per unit frequency range at frequency ω_k. The quantities X_k, Y_k, U_k, and V_k are independently distributed Gaussian random variables such that

$$\begin{aligned}&\langle X_k\rangle_{Av}=\langle Y_k\rangle_{Av}=\langle U_k\rangle_{Av}=\langle V_k\rangle_{Av}=0,\\&\langle X_k^2\rangle_{Av}=\langle Y_k^2\rangle_{Av}=\langle U_k^2\rangle_{Av}=\langle V_k^2\rangle_{Av}=1.\end{aligned}\tag{6}$$

Thus the probability that a particular one of these variables has a value between x and $x+dx$ is

$$P_G(x)dx=[1/(2\pi)^{\frac{1}{2}}]\exp(-x^2/2)dx.\tag{7}$$

The representation given in Eqs. (5) and (7) is strictly correct only in the limit that the number of pulses per unit time becomes infinite. However it is a sufficiently good approximation if the number of pulses emitted in a time equal to the reciprocal line breadth is large compared to one.

Equation (5) can be immediately written in the form of Eq. (1) by using the definitions

$$\begin{aligned}X(t)&=\sum p_k[X_k\cos(\omega_k-\omega_0)t\\&\qquad+U_k\sin(\omega_k-\omega_0)t],\\U(t)&=\sum p_k[-X_k\sin(\omega_k-\omega_0)t\\&\qquad+U_k\cos(\omega_k-\omega_0)t],\\Y(t)&=\sum p_k[Y_k\cos(\omega_k-\omega_0)t\\&\qquad+V_k\sin(\omega_k-\omega_0)t],\\V(t)&=\sum p_k[-Y_k\sin(\omega_k-\omega_0)t\\&\qquad+V_k\cos(\omega_k-\omega_0)t].\end{aligned}\tag{8}$$

The quantities $X(t)$, $Y(t)$, $U(t)$, and $V(t)$, being linear combinations of Gaussianly distributed variables, are themselves Gaussianly distributed. Furthermore they are independently distributed because of the orthogonality property of the coefficients in their expansions in terms of X_k, Y_k, U_k, and V_k. The quantity p appearing in Eq. (2) is given by

$$p^2=\sum_{k=0}^{\infty}p_k^2.\tag{9}$$

Thus p^2 is $2\pi/c$ times the average power flow per unit area in the beam. The only remaining property of the functions $X(t)$, $Y(t)$, $U(t)$, and $V(t)$ to be proved is that they are slowly varying compared to $\cos\omega_0 t$ and $\sin\omega_0 t$. But this follows immediately from Eq. (8) and the fact that p_k is appreciable only if ω_k is close to ω_0.

Our second method of obtaining the fundamental representation given in Eqs. (1) to (3) is similar to the derivation of the Maxwellian distribution of velocities in a gas from the assumption that the Cartesian components of the velocities are independently distributed. The analogous assumption which we shall make is that E_x and E_y are independently distributed. More specifically we shall assume that the functions $X(t)$ and $U(t)$ in Eq. (1) are distributed independently of the functions $Y(t)$ and $V(t)$.[3] This independence property followed directly from the model assumed in the above derivation. However it is more general than the model there assumed and so may logically be taken as the starting point in our second derivation. In this derivation we begin by arbitrarily writing the field components in the form given in Eq. (1). Since the electric field is assumed to have appreciably large Fourier components only for angular frequencies close to ω_0 it follows immediately that the functions $X(t)$, $Y(t)$, $U(t)$, and $V(t)$ are slowly varying compared to $\cos\omega_0 t$ and $\sin\omega_0 t$. It remains to be proved that these four functions are Gaussianly distributed and that $X(t)$ is independent of $U(t)$ and $Y(t)$ is independent of $V(t)$. In addition to the fundamental assumption that $X(t)$ and $U(t)$ are independent of $Y(t)$ and $V(t)$ we shall assume that the statistical properties of the field components are independent of the orientation of the x and y axes. This means, for example, that, for any value of θ the quantity $X(t)\cos\theta+Y(t)\sin\theta$ must have the same distribution function as the quantity $X(t)$. The fact that $X(t)$, $Y(t)$, $U(t)$, and $V(t)$ are Gaussianly distributed is then an immediate consequence of the following theorem:

[3] The assumption that E_x and E_y are statistically independent can be justified in terms of the somewhat more physical assumption that the statistical properties of an intense beam of unpolarized light cannot be changed by passing the light through a wave plate. By the use of a wave plate of suitable thickness, one field component can be retarded with respect to the other by any desired amount. It is evident that no possible type of correlation between E_x and E_y could remain invariant under this general class of transformations.

If A and B are independently distributed random variables and for every θ the variable $A\cos\theta+B\sin\theta$ has the same distribution as A (or B), then A and B are Gaussianly distributed. This theorem was proved in utmost generality by M. Kac and H. Steinhaus.[4] Their proof proceeds by obtaining a simple functional equation for the characteristic function of the distribution. Since A and $A\cos\theta+B\sin\theta$ have the same distribution it follows that for every real ξ the average value of $e^{i\xi A}$ satisfies the relation

$$f(\xi)=\langle e^{i\xi A}\rangle_{Av}=\langle e^{i\xi(A\cos\theta+B\sin\theta)}\rangle_{Av}. \quad (10)$$

But since A and B are independently distributed Eq. (10) can be rewritten

$$f(\xi)=\langle e^{i\xi A\cos\theta}\rangle_{Av}\langle e^{i\xi B\sin\theta}\rangle_{Av}$$

$$=f(\xi\cos\theta)f(\xi\sin\theta); \quad (11)$$

or

$$f((x^2+y^2)^{\frac{1}{2}})=f(x)f(y), \quad (12)$$

where

$$x=\xi\cos\theta; \quad y=\xi\sin\theta. \quad (13)$$

It is easy to show that the only continuous solution of this well known equation is of the form

$$f(\xi)=\exp(\lambda\xi^2). \quad (14)$$

Furthermore λ is negative since

$$|f(\xi)|=|\langle e^{i\xi A}\rangle_{Av}|\leqslant 1. \quad (15)$$

Thus, since $f(\xi)$ is the characteristic function of a Gaussian distribution the theorem is proved. In our application of this theorem the random variable A corresponds to $X(t)$ (or $U(t)$) and the random variable B corresponds to $Y(t)$ (or $V(t)$).

The proof that $X(t)$ is independent of $U(t)$ and $Y(t)$ is independent of $V(t)$ is based on the fact that $X(t)$, $Y(t)$, $U(t)$, and $V(t)$ are Gaussianly distributed, and the additional assumption that $A(t)$ is independent of $\delta_x(t)$ and $B(t)$ is independent of $\delta_y(t)$. This latter assumption is clearly justified since $\delta_x(t)$ and $\delta_y(t)$ depend on the origin of time which is arbitrary. Let $P_A(a)da$ be the probability that $A(t)$ lies between a and $a+da$. Then, since $A(t)$ and $\delta_x(t)$ are independent, and in view of the relations given in Eq. (1), the probability that $X(t)$ lies between x and $x+dx$, and simultaneously $U(t)$ lies between u and $u+du$ is

$$P_{X,U}(x,u)dxdu=\frac{P_A((x^2+u^2)^{\frac{1}{2}})}{2\pi(x^2+u^2)^{\frac{1}{2}}}dxdu$$

$$=g(x^2+u^2)dxdu. \quad (16)$$

Hence

$$P_X(x)=\frac{1}{p(2\pi)^{\frac{1}{2}}}\exp\left(-\frac{x^2}{2p^2}\right)$$

$$=\int_{-\infty}^{+\infty}P_{X,U}(x,u)du$$

$$=\int_{-\infty}^{+\infty}g(x^2+u^2)du. \quad (17)$$

Replacing x^2 by x^2+z^2 we have

$$\frac{1}{p(2\pi)^{\frac{1}{2}}}\exp\left(-\frac{(x^2+z^2)}{2p^2}\right)$$

$$=\int_{-\infty}^{+\infty}g(x^2+z^2+u^2)du; \quad (18)$$

so that

$$\frac{1}{p(2\pi)^{\frac{1}{2}}}\exp\left(-\frac{x^2}{2p^2}\right)\int_{-\infty}^{+\infty}\exp\left(-\frac{z^2}{2p^2}\right)dz$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}g(x^2+z^2+u^2)dudz, \quad (19)$$

or

$$\exp\left(-\frac{x^2}{2p^2}\right)=\pi\int_{x^2}^{\infty}g(y)dy. \quad (20)$$

By differentiating both sides of the above equation with respect to x^2 we find

$$\frac{1}{2p^2}\exp\left(-\frac{x^2}{2p^2}\right)=\pi g(x^2), \quad (21)$$

so that

$$P_{X,U}(x,u)dxdu$$

$$=\frac{1}{2\pi p^2}\exp\left(-\frac{(x^2+u^2)}{2p^2}\right)dxdu. \quad (22)$$

Since the joint distribution function of $X(t)$ and $U(t)$ is the product of the individual distribution functions the variables $X(t)$ and $U(t)$ are independent. This completes our second derivation of Eqs. (1) to (3).

[4] M. Kac and H. Steinhaus, Studia Mathematica **6**, 89 (1936).

The difference between polarized and unpolarized light can be easily understood in terms of our fundamental representation as given in Eqs. (1) to (3). For light which is elliptically polarized in the usual sense the quantities $A(t)$, $B(t)$, $\delta_x(t)$, and $\delta_y(t)$ still are functions of time, but they vary in such a way that $A(t)/B(t)$ and $\delta_x(t)-\delta_y(t)$ remain constant. Hence though the amplitude and phase of the ellipse traced by the electric vector vary with time, the eccentricity and orientation remain constant. This can be accomplished in the laboratory by first removing one of the components of the electric field by a Nicol prism or Polaroid disk and then passing the resultant plane polarized beam through a wave plate. Clearly variations in the phase and amplitude of the plane polarized beam cannot cause variations in phase differences and amplitude ratios of components of the field after the wave plate.

An interesting paradox is encountered if one considers the fluctuations in the energy flowing past a point in the beam in a time interval τ. Let

$$W_\tau(t)=\int_t^{t+\tau}(A^2(t')+B^2(t'))dt'. \qquad (23)$$

The statistical behavior of the quantity $W_\tau(t)$ has been carefully investigated in another connection by M. Kac. His calculations are based on an equation corresponding to Eq. (5). It is easy to show that the quantity

$$\Delta=\frac{\langle W_\tau^2(t)\rangle_{Av}-\langle W_\tau(t)\rangle_{Av}^2}{\langle W_\tau(t)\rangle_{Av}^2}$$

depends only on τ and not on $\langle W_\tau(t)\rangle_{Av}$. Its limiting value for small τ is $\frac{1}{2}$ while for large τ it approaches zero. For values of τ of order of magnitude of the reciprocal line breadth the value of Δ depends on the line shape. According to quantum theory the quantity $W_\tau(t)$ is proportional to the number of photons incident in the time interval τ. One would expect this number n to have a Poisson distribution which would mean that Δ should be inversely proportional to n. As was pointed out by V. Weisskopf, the explanation of this paradox lies in the fact that the photons obey Einstein-Bose statistics. The photons therefore have a greater tendency for bunching than they would have if they were independent.[5] Hence the fluctuations in $W_\tau(t)$ are increased so that if the intensity is sufficiently large Δ is independent of $\langle W_\tau(t)\rangle_{Av}$ instead of being inversely proportional to it.

STATISTICAL PROPERTIES OF THE INSTANTANEOUS ELLIPSE

The distribution function of the parameters of the instantaneous ellipse can be calculated quite easily from Eqs. (1) to (3). Let $M(t)$ and $N(t)$ be respectively the semi-major and semi-minor axes of the instantaneous ellipse. Then[6]

$$M^2(t)+N^2(t)=A^2(t)+B^2(t)=R^2(t),$$

$$S(t)=\frac{2M(t)N(t)}{R^2(t)}=\frac{2A(t)B(t)}{R^2(t)}|\sin\delta(t)|, \qquad (24)$$

$$\delta(t)=\delta_x(t)-\delta_y(t).$$

Let

$$\Phi(t)=\tan^{-1}\frac{B(t)}{A(t)} \quad 0\leqq\Phi(t)\leqq\frac{\pi}{2}. \qquad (25)$$

Then the probability that $R(t)$ lies between r and $r+dr$ and simultaneously $\Phi(t)$ lies between φ and $\varphi+d\varphi$ is

$$P_{R,\Phi}(r,\varphi)drd\varphi=P_A(r\cos\varphi)P_B(r\sin\varphi)rdrd\varphi$$

$$=\frac{r^3}{2p^4}\sin 2\varphi\exp\left(-\frac{r^2}{2p^2}\right)drd\varphi. \qquad (26)$$

The quantity $rdrd\varphi$ represents the element of area in the A-B plane, and we have used Eq. (3) for $P_A(x)$ and $P_B(x)$. From Eq. (24)

$$S(t)=\sin 2\Phi(t)\,|\sin\delta(t)|. \qquad (27)$$

Since $\delta(t)$ is uniformly distributed between zero and 2π the probability that $|\sin\delta(t)|$ lies between σ and $\sigma+d\sigma$ is

$$P_{|\sin\delta|}(\sigma)d\sigma=\frac{2}{\pi}d\sin^{-1}\sigma=\frac{2}{\pi}\frac{d\sigma}{(1-\sigma^2)^{\frac{1}{2}}}. \qquad (28)$$

If the values of R and Φ are fixed, the probability that S lies between s and $s+ds$, which is obtained by substituting $ds/\sin 2\Phi$ for $d\sigma$ and $s/\sin 2\Phi$ for

[5] L. Brillouin, *Die Quantenstatistik* (Julius Springer, Berlin, 1931), p. 111, 177.
[6] M. Born, *Optik* (Julius Springer, Berlin, 1933), p. 23.

σ in Eq. (28), is

$$P_{|\sin\delta|}\left(\frac{s}{\sin 2\Phi}\right)\frac{ds}{\sin 2\Phi} = \frac{2}{\pi}\frac{ds}{\sin 2\Phi\left(1-\frac{s^2}{\sin^2 2\Phi}\right)^{\frac{1}{2}}}. \quad (29)$$

Hence the probability that R lies between r and $r+dr$ while S lies between s and $s+ds$ is

$$P_{S,R}(s,r)dsdr = dsdr\int \frac{2}{\pi}\frac{P_{R,\Phi}(r,\varphi)d\varphi}{\sin^2\varphi\left(1-\frac{s^2}{\sin^2 2\varphi}\right)^{\frac{1}{2}}}. \quad (30)$$

The integral is to be extended over that part of the first quadrant for which $\sin 2\varphi$ is greater than s. From Eq. (26)

$$P_{S,R}(s,r)dsdr = dsdr\frac{r^3}{\pi p^2}\exp\left(-\frac{r^2}{2p^2}\right) \times\int \frac{d\varphi}{\left(1-\frac{s^2}{\sin^2 2\varphi}\right)^{\frac{1}{2}}}. \quad (31)$$

But

$$\int \frac{d\varphi}{\left(1-\frac{s^2}{\sin^2 2\varphi}\right)^{\frac{1}{2}}} = -\frac{1}{2}\int\frac{dy}{(1-s^2-y^2)^{\frac{1}{2}}} = \frac{\pi}{2} \quad (32)$$

$$y=\cos 2\varphi$$

since the range of integration is clearly from $y=(1-s^2)^{\frac{1}{2}}$ to $y=-(1-s^2)^{\frac{1}{2}}$. Thus

$$P_{S,R}(s,r)dsdr = \frac{r^3}{2p^4}\exp\left(-\frac{r^2}{2p^2}\right)dsdr. \quad (33)$$

Equation (33) shows that S is uniformly distributed between zero and one, and that R is distributed independently of S according to the law

$$P_R(r)dr = \frac{r^3}{2p^4}\exp(-r^2/2p^2)dr. \quad (34)$$

The distribution of the ratio

$$L(t)=N(t)/M(t) \quad (35)$$

can be found immediately from Eq. (33). We have

$$S=2L/(1+L^2), \quad (36)$$

so that

$$dS=2(1-L^2)dL/(1+L^2)^2. \quad (37)$$

Hence the joint distribution function of $L(t)$ and $R(t)$ is

$$P_{L,R}(l,r)dldr = \frac{(1-l^2)}{(1+l^2)^2}\frac{r^3}{p^4} \times\exp\left(-\frac{r^2}{2p^2}\right)dldr, \quad (38)$$

and the distribution function of $L(t)$ is

$$P_L(l)dl = [2(1-l^2)/(1+l^2)^2]dl. \quad (39)$$

The function $P_L(l)$ has its maximum value for l equal to zero, and decreases monotonically to zero when l equals one. The median value of $L(t)$ is related to the median value of $S(t)$ by Eq. (36). Since the median value of $S(t)$ is $\frac{1}{2}$ we find for the median value of $L(t)$

$$L_m=.268. \quad (40)$$

We therefore have the rather surprising result that more than half the time, the ellipse being traced out by the electric vector has a major axis more than three and one half times as large as the minor axis. The quantities $L(t)$ and $M(t)$ are not independently distributed but have the joint distribution function

$$P_{L,M}(l,m)dldm = (1-l^2)\frac{m^3}{p^4}\exp\left[-\frac{m^2(1+l^2)}{2p^2}\right]dldm. \quad (41)$$

The quantities $M(t)$ and $N(t)$ have the joint distribution function

$$P_{M,N}(m,n)dmdn = \frac{(m^2-n^2)}{p^4}\exp\left[-\frac{(m^2+n^2)}{2p^2}\right]dmdn. \quad (42)$$

It is interesting to note that the distribution function $P_S(s)$ of $S(t)$ can be derived directly

from the requirement that it be unchanged if the light is passed through a quarter wave plate. This condition leads to an integral equation for $P_S(s)$ which has a constant as its only solution. An alternate and elegant derivation of the same result using Poincaré's representation for polarized light has been suggested by H. Mueller and L. Tisza.[7]

CONCLUDING REMARKS

The statistical properties of unpolarized light which are discussed above cannot be measured with ordinary optical equipment such as wave plates and Nicol prisms. The standard experiments cannot be performed in sufficiently short times and so give only the average values of the field components. However the difficulty is only one of experimental technique for the parameters of the instantaneous ellipse are measurable in principle if the beam is sufficiently intense. In fact an actual experiment might be possible if radiation in the radio or microwave region were used. Although a spectral line in the microwave region could conceivably be employed, it would probably be more convenient to use a narrow spectral range of high temperature black body radiation. The experimental procedure would be to employ two narrow band receivers fed by antennas responsive to electric fields in perpendicular directions. The output of the receivers (possibly after beating with a standard frequency) would be connected to the two pairs of deflection plates of an oscilloscope. The pattern on the oscilloscope screen would then correspond to the pattern traced out by that part of the incident electric field in the narrow frequency range to which the receivers were responsive. (The fact that the narrow frequency range is singled out by filters in the receivers rather than by a filter acting on the radiation before it reaches the antennas does not alter the final results.)

An interesting feature of the experiment just described is that if the receiver noise were large enough to completely overshadow the signal due to the incoming radiation, the statistical behavior of the pattern on the oscilloscope screen would be unchanged. The reason is that the most important forms of receiver noise are fundamentally due to shot effect and therefore have the statistical properties corresponding to Eqs. (5) and (7). This suggests that a simple way to obtain a model of the behavior of the electric vector in natural monochromatic light is to feed two narrow band amplifiers with shot noise from two diodes and put their outputs on the deflection plates of an oscilloscope. The electron beam will then trace out an ellipse whose parameters fluctuate at a rate inversely proportional to the band width and have the statistical properties described in the preceding section.

[7] Private communication.

Ultraviolet Absorption Method for the Determination of Polyunsaturated Constituents in Fatty Materials*

BROOKS A. BRICE AND MARGARET L. SWAIN
*Eastern Regional Research Laboratory,** Philadelphia, Pennsylvania*

INTRODUCTION

VEGETABLE oils and animal fats are complex mixtures consisting essentially of saturated and unsaturated mixed fatty acid glycerides. The distribution of fatty acid components in different fatty materials varies considerably, and the complete analysis for these components in a given material poses a complex problem. The fatty acid constituents of unsaturation higher than oleic acid (1 double bond) likely to be present are linoleic (2 double bonds), linolenic (3 double bonds), and arachidonic acids (4 double bonds), and their conjugated isomers. Because of the high reactivity of these constituents, their presence even in minor and trace proportions may have considerable significance in research, industrial, and food uses of fats and oils.

The chemical method of analysis for the non-conjugated polyunsaturated constituents, linoleic, linolenic, and arachidonic acids, based on determination of iodine number and thiocyanogen number, is time-consuming and is inaccurate when these compounds are present in small proportions. Supplementary insoluble bromide tests are likewise insensitive and unreliable in such cases. No reliable chemical method exists for determining the minor conjugated constituents of oils and fats.

Recently proposed spectrophotometric methods offer the possibility of furnishing a method of analysis for these polyunsaturated constituents which is superior to chemical methods in speed, sensitivity, and specificity. Direct ultraviolet spectrophotometric analysis for the conjugated components is relatively simple, involving difficulties only when they are present in small proportions. Ultraviolet spectrophotometric analysis for the non-conjugated components, which have no absorption maxima above 200 mμ, may be accomplished by converting them to their conjugated isomers by alkali-isomerization followed by ultraviolet absorption measurements, as shown by Mitchell, Kraybill, and Zscheile,[1] and by Beadle and Kraybill.[2] This method also involves difficulties when only small proportions of these constituents are present.

The purpose of the present investigation was to combine, modify, and improve existing spectrophotometric methods for application principally to materials having small proportions of polyunsaturated constituents, such as animal fats and their soaps, partially hydrogenated fats, and purified fatty preparations.

The modified method proposed here comprises measurement of the ultraviolet absorption of a sample before and after isomerization, the latter in an improved medium; correction of data for extraneous absorption; correction of data for conjugated constituents originally present and remaining undestroyed by the isomerization treatment; and calculation of the proportions of conjugated and non-conjugated diene, triene, and tetraene fatty acid constituents in the sample.

Throughout this paper, specific extinction coefficient, k, will be defined as

$$k = D/bc = (-\log T)/bc,$$

where D is the spectral density and T the transmittance of a solution relative to that of the solvent in an equal cell, b is the inside length (in centimeters) of the cell used, and c is the concentration of the solution in grams per liter of solution. All data were obtained with a Beckman Model DU photoelectric spectrophotometer, using a hydrogen lamp and an absorption cell compartment for cells up to 10 cm long.

CONJUGATED CONSTITUENTS

Bradley and Richardson[3] have analyzed certain vegetable oils for conjugated fatty acids

* Presented before the Optical Society of America, New York, New York, October 20–21, 1944. J. Opt. Soc. Am. **34**, 772 (1944) (Abstract).

** One of the laboratories of the Bureau of Agricultural and Industrial Chemistry, Agricultural Research Administration, United States Department of Agriculture.

[1] J. H. Mitchell, Jr., H. R. Kraybill, and F. P. Zscheile, Ind. Eng. Chem. Anal. Ed. **15**, 1 (1943).

[2] B. W. Beadle and H. R. Kraybill, J. Am. Chem. Soc. **66**, 1232 (1944).

[3] T. F. Bradley and D. Richardson, Ind. Eng. Chem. **34**, 237 (1942).

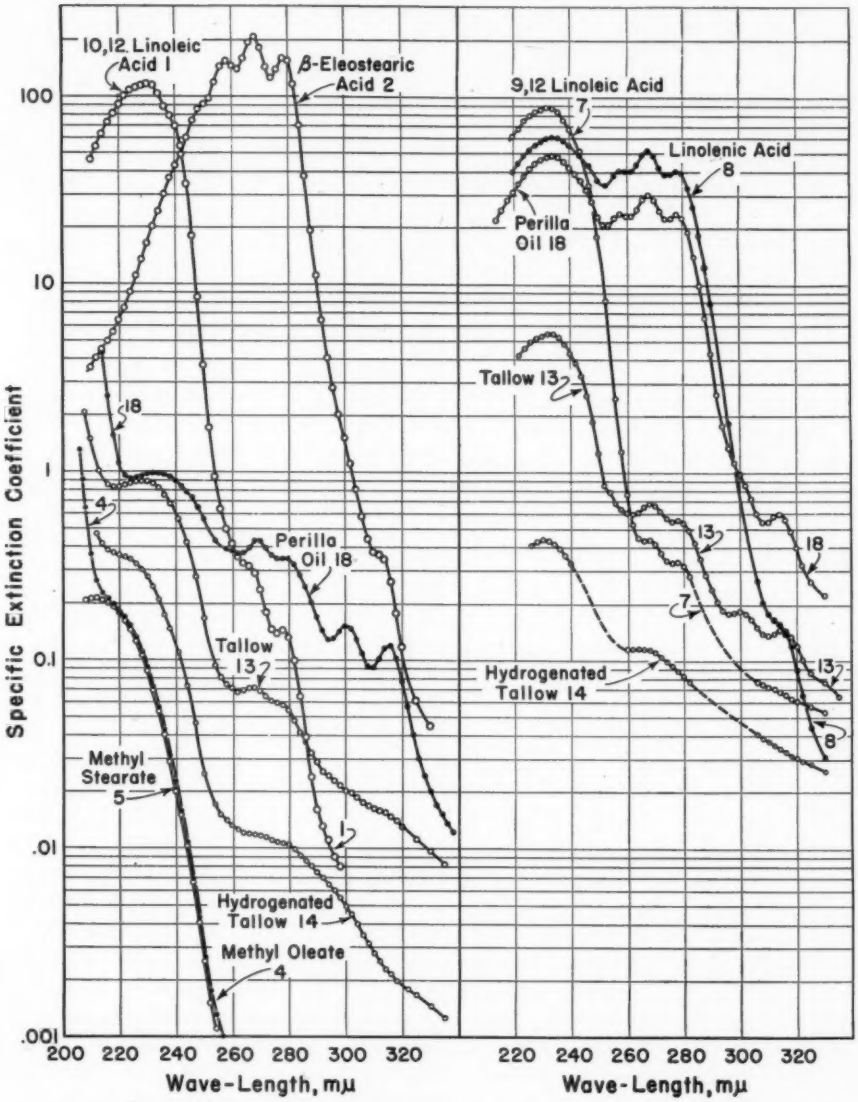


FIG. 1. Ultraviolet absorption curves for ethanol solutions of purified β-eleostearic acid and 10,12 linoleic acid standards; neohexane solutions of purified methyl oleate and methyl stearate, and typical oil and fat samples containing small proportions of conjugated fatty acid glycerides. Slit widths 1.5 to 2 mμ.

FIG. 2. Ultraviolet absorption curves for ethanol solutions of fatty materials after isomerization in KOH-glycerol solution for 30 minutes at 180°C: purified 9,12 linoleic acid and linolenic acid standards; and the typical oil and fat samples shown in Fig. 1. Slit widths approximately 3.5, 2.5 and 4.5 mμ at 232, 268, and 316 mμ, respectively.

TABLE I. Ultraviolet absorption data for purified fatty acids and esters: observed or reported values for wave-length λ_{max} and specific extinction coefficient k at absorption maximum used; values tentatively adopted as standard in analysis for conjugated fatty acid constituents; and ratio of k value at λ_{max} to average value of k at $\lambda_{max}+6$ mμ and $\lambda_{max}-6$ mμ, i.e., $k/\Delta k$.

Sample No.	Double bonds	Compound	Solvent	Observed			Used in analysis		
				mμ λ_{max}	k	$k/\Delta k$	mμ λ	k	$k/\Delta k$
1	2[a]	10,12 linoleic acid	Neohexane	231	119	—	232	119	—
			Ethanol	231	117	—			
2	3[a]	β-eleostearic acid	Neohexane	267	214	2.71	268	214	2.8
			Ethanol	268	209	2.78			
—	4[a]	Parinaric acid[b]	Ethanol	320	203	1.7	316	220	2.5
—	4[a]	Tetraene acid from isomerized arachidonic acid	Ethanol	315[c]	220[c]	2.5[d]			
3	1	Methyl oleate	Ethanol	232[e]	.077	—			
4	1	Methyl oleate	Neohexane	232[e]	.072	—	232	.07	
4-A	1	Oleic acid	Neohexane	232[e]	.032	—	232	.03	
5	0	Methyl stearate	Neohexane	232[e]	.070	—	232	.07	
6	0	Stearic acid	Neohexane	232[e]	.025	—	232	.03	

[a] Conjugated.
[b] Data obtained by comparator measurements on Fig. 1 of reference 6.
[c] From reference 12.
[d] See footnote 13.
[e] No maxima or inflections observed at this wave-length.

simply by calculating the ratios of the specific extinction coefficients for the oils to those for purified 9,11 linoleic acid,[4] α-eleostearic acid,[5] and parinaric acid,[6] at the wave-lengths of the absorption maxima of the acids, assuming that all the observed absorption for the oils at these wave-lengths was due to the conjugated fatty constituents. Brode, Patterson, Brown, and Frankel[7] have attempted to take into account the overlapping of the absorption curves of the diene, triene, and tetraene conjugated acids, and have set up simultaneous equations for calculating their proportions in mixtures. Kass[8] has applied a similar procedure to the analysis of drying oils, using more recent absorption data for purified conjugated acids. This method is particularly applicable to materials having substantial proportions of conjugated fatty acids and relatively free of interfering absorption by extraneous compounds.

The methods mentioned above do not yield accurate analyses for materials such as animal fats, many vegetable oils, and many fatty preparations, which ordinarily contain, in decreasing order of abundance, diene, triene, and tetraene conjugated fatty acid compounds as minor and trace constituents. Such materials almost invariably contain other absorbing bodies, consisting probably of oxidation and polymerization products, pigments, and other unknown compounds.

Ultraviolet absorption curves for typical samples of perilla oil, tallow, and a partially hydrogenated tallow before isomerization are shown in Fig. 1, and after isomerization in Fig. 2. Overlapping of absorption by the various fatty acid constituents, and the presence in some cases of absorption by extraneous compounds are apparent. For the partially hydrogenated tallow 14, in both Fig. 1 and Fig. 2, the absence of absorption maxima or inflections in the regions of absorption by conjugated triene acids (near 268 and 278 mμ) and by conjugated tetraene acids (near 301 and 316 mμ) indicates with a high degree of sensitivity absence of triene and

[4] L. J. van der Hulst, Rec. Trav. Chim. des Pays-Bas **54**, 639, 644 (1935).
[5] A. Dingwall and J. C. Thomson, J. Am. Chem. Soc. **56**, 899 (1934).
[6] H. P. Kaufmann, J. Baltes, and S. Funke, Fette u. Seifen **45**, 302 (1938).
[7] W. R. Brode, J. W. Patterson, J. B. Brown, and J. Frankel, Ind. Eng. Chem. Anal. Ed. **16**, 77 (1944).
[8] J. P. Kass, *Protective and Decorative Coatings*, edited by J. J. Matiello (John Wiley and Sons, New York, 1944), Chap. 12.

tetraene fatty acid constituents, although appreciable absorption is observed in these regions. A weak, broad, smooth absorption band near 280 mμ has been observed with many selectively hydrogenated tallows and their soaps. A similar broad, smooth absorption band, but with a maximum near 264 mμ, has been observed in alkali-isomerized samples of partially hydrogenated lards and tallows and their soaps, and methyl oleate. While the origin of these bands near 280 and 264 mμ is at present obscure, they cannot be attributed to conjugated triene fatty acid components.

Ultraviolet absorption curves are shown in Fig. 1 for purified β-eleostearic acid, 10,12 linoleic acid, methyl oleate, and methyl stearate. Data for these and other purified fatty acids are summarized in Table I.

The specific extinction coefficient 119 at 231 mμ obtained for the 10,12 linoleic acid sample (prepared from dehydrated castor oil) in neohexane is in satisfactory agreement with published values[9] for 9,11 and 10,12 linoleic acid. The shape of the curve in the region 260–280 mμ indicates the presence of a small impurity of conjugated triene acid, probably less than 0.1 percent. A wave-length of 232 mμ was adopted for analytical purposes in order to coordinate data with analyses involving alkali isomerization.

The specific extinction coefficient 214 at 267 mμ obtained for the β-eleostearic acid (prepared from tung oil) in neohexane is in satisfactory agreement with published values,[10] and was adopted as a standard for use in analysis for conjugated triene acids. The weak maximum near 313 mμ in Fig. 1 for this sample indicates the presence of a small impurity of conjugated tetraene acid. The higher value shown for neohexane solutions as compared with alcohol applies only to freshly prepared solutions.[11] A wave-length of 268 mμ was adopted for analytical purposes, since maxima in solutions of samples are usually at 268 mμ.

The only data available to use as standards in the analysis for tetraene conjugated acids are those for parinaric acid,[6] a naturally occurring C_{18} acid, and those for a C_{20} conjugated tetraene fatty acid isolated from isomerized arachidonic acid.[12] For the former acid, absorption maxima at 292, 307, and 320 mμ were reported, with a k value of about 203 at the 320 mμ maximum. For the latter acid, maxima were reported at 286, 300, and 315 mμ, with a k value of 220 at 315 mμ. Since conjugated tetraene acids are usually present only in trace amounts in vegetable oils and animal fats, it matters little which of these acids is chosen as a standard. However, until further work is done on such acids, we have preferred to use the data on the conjugated tetraene acid from isomerized arachidonic acid as a standard for analysis, since we have observed an absorption maximum near 316 mμ and an inflection near 301 mμ in many fatty materials.

Examination of the absorption curves of Fig. 1 and a large number of curves for vegetable oils and animal fats leads to the following conclusions: (*a*) reliable analyses for conjugated diene components will in many cases necessitate correcting the observed absorption at 232 mμ for absorption by the carboxyl radical (see Fig. 1, curve for methyl stearate 5 and methyl oleate 4) and for multiple C=C groups (see steep rise in curve below 220 mμ for perilla oil 18, which is rich in linolenic acid; also note in Table I that the single C=C group apparently has no appreciable absorption at 232 mμ); (*b*) reliable analyses for conjugated triene components will involve correction of observed absorption near 268 mμ for background absorption by conjugated diene and unknown constituents; (*c*) reliable analyses for conjugated tetraene components will involve correction of observed absorption near 316 mμ for background absorption by conjugated triene and unknown constituents. The same considerations apply also to the examination of curves such as

[9] Kass (reference 8) gives 115 at 232 mμ for both 9,11 and 10,12 linoleic acids; Kerns, Belkengren, Clark, and Miller, J. Opt. Soc. Am. **31**, 271 (1941) report 116 at 232 mμ for 10,12 linoleic acid; our measurements with a comparator on van der Hulst's curves (reference 4, Fig. 3, page 641 and Fig. 6, page 649) indicate values of 119 and 126 for 9,11 linoleic acid in hexane.

[10] Kass (reference 8) reports a value of 215 at 268 mμ for both β- and pseudo-eleostearic acids; Dingwall and Thomson (reference 5) about 189 for β-eleostearic acid in ethanol; our measurements with a comparator on van der Hulst's curve (reference 4, Fig. 5, page 642) indicates about 194 for β-eleostearic acid in hexane.

[11] Instability of preparations of α-eleostearic acid have been discussed by R. T. O'Connor, D. C. Heinzelman, A. F. Freeman, and F. C. Pack, Ind. Eng. Chem. Anal. Ed. (in press).

[12] D. T. Mowry, W. R. Brode, and J. B. Brown, J. Biol. Chem. **142**, 671 (1942).

those of Fig. 2 in the analysis for the non-conjugated fatty acids by the method of alkali-isomerization.

In the equations which follow, subscripts 2, 3, and 4 refer to the number of C=C bonds, and subscripts 232, 268, etc., refer to wave-lengths. The specific extinction coefficient due to conjugated diene fatty acid components will be approximately

$$k_2 = k_{232} - 0.07 - \cdots \text{ for fats and esters,} \quad (1)$$

$$k_2 = k_{232} - 0.03 - \cdots \text{ for soaps and acids,} \quad (1a)$$

where k_{232} is the observed coefficient at 232 mμ; 0.07 and 0.03 are the specific extinction coefficients at 232 mμ for purified methyl stearate and stearic acid, respectively (see Fig. 1 and Table I). Evidence that the value 0.07 is applicable to glycerides as well as to monohydric esters exists in the observation that many hydrogenated tallows having only traces of polyunsaturated constituents show k values of about 0.09 at 232 mμ. The equations as written are applicable principally to materials having small proportions of non-conjugated polyunsaturated constituents. For materials containing substantial proportions of linoleic and linolenic acids, appropriate terms will have to be added to these equations as reliable data become available.

An approximate correction for overlapping and background absorption in the regions of absorption by the conjugated triene and tetraene fatty acid compounds may be accomplished by determining the prominence of the absorption maxima near 268 and 316 mμ in terms of specific extinction coefficient differences at and near the maxima.

It can be shown that for a solution of a mixture containing β-eleostearic acid and one or more absorbing constituents having linear absorption curves between 262 and 274 mμ, the specific extinction coefficient due to the β-eleostearic acid in the mixture is

$$k_3 = 2.8[k_{268} - \tfrac{1}{2}(k_{262} + k_{274})], \quad (2)$$

where k_{268}, k_{262}, and k_{274} are observed coefficients for the mixture at the indicated wave-lengths, and 2.8 is a constant evaluated from measurements on solutions of pure β-eleostearic acid.

The derivation and application of this relation is illustrated in Fig. 3. For the pure acid in alcohol solution, the ratio $k_{268}/[k_{268} - \frac{1}{2}(k_{262} + k_{274})] = k/\Delta k = 209/75.3 = 2.8$. In a solution of β-eleostearic acid of unknown concentration, the value of Δk will be proportional to the concentration and the constant 2.8 will be independent of the concentration. In a mixture containing the acid and absorbing constituents having linear absorption between 262 and 274 mμ, the contribution of the latter to Δk will be zero, and thus the concentration of β-eleostearic acid in the mixture is still proportional to Δk. Hence the expression for k_3 in Eq. (2).

Applying Eq. (2) to curve *B* in Fig. 3 for a typical tallow, for which the observed values of k_{268}, k_{262}, and k_{274} are 0.0542, 0.0508, and 0.0434, respectively, we calculate $\Delta k = 0.0071$ and $k_3 = 0.0200$, the coefficient due to eleostearic acid. This leaves a value of 0.0342 due presumably to extraneous constituents. Verification of this result is illustrated by curve *C*, which was obtained by subtracting from curve *B* increasing fractions of values of curve *A* until no maxima or minima remained in the resulting curve. The value of k at 268 mμ in curve *C*, representing tallow 14 minus 0.009 percent β-eleostearic acid, was 0.0353, which is in good agreement with the value 0.0342 calculated by use of Eq. (2) for the background absorption. Thus in this example, less than half the observed absorption at 268 mμ is due to conjugated triene fatty acid constituents.

It should be pointed out that the value of the constant 2.8 in Eq. (2) varies somewhat for the different isomers of eleostearic acid, for different solvents (Table I), and for different slit widths. However, for β- or pseudo-eleostearic acids, in neohexane or alcohol, and with spectrum band widths less than 2 mμ, any deviations in this value are within about 4 percent and thus give rise to much smaller errors than would be made by neglecting background correction.

Similar treatment in the region of the longest wave-length absorption band of conjugated tetraene acids leads to the following relation for the specific extinction coefficient of conjugated tetraene acids in the presence of other constituents having linear absorption curves between 310 and 322 mμ:

$$k_4 = 2.5[k_{316} - \tfrac{1}{2}(k_{310} + k_{322})], \quad (3)$$

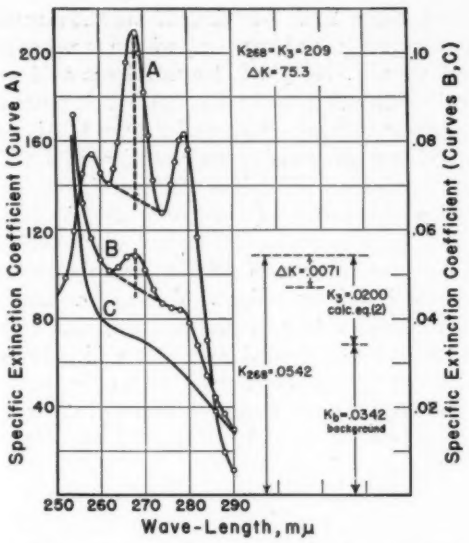


FIG. 3. Absorption curves: *A*, purified β-eleostearic acid 2 in ethanol; *B*, tallow 16 in neohexane; *C*, calculated curve of "background" absorption, i.e., tallow 16 minus 0.009 percent β-eleostearic acid.

where k_{316}, k_{310}, and k_{322} are the observed coefficients of the mixture, and 2.5 is a constant evaluated from isomerized arachidonic acid.[13] Background absorption in this region is in general more nearly linear than is the case in the conjugated triene region.

The concentrations of conjugated diene, triene, and tetraene constituents, expressed as percent of fatty acids, present in a given sample are calculated by the following equations:

$$C_2 = 100k_2/119, \quad (4)$$

$$C_3 = 100k_3/214, \quad (5)$$

$$C_4 = 100k_4/220, \quad (6)$$

where k_2, k_3, and k_4 are the corrected specific extinction coefficients calculated from Eqs. (1) or (1a)–(3), and the constants in the denominators are the coefficients for the standards in Table I.

Results on the analysis of typical samples for conjugated constituents appear in Tables V and VI in conjunction with other data.

[13] In the absence of detailed data on the conjugated tetraene acid isolated from isomerized arachidonic acid (reference 12), this constant was tentatively evaluated from data of Beadle and Kraybill (reference 2) for isomerized arachidonic acid. The authors are indebted to B. W. Beadle for the necessary data for this calculation.

NON-CONJUGATED CONSTITUENTS

The method of Mitchell, Kraybill, and Zscheile[1] for the analysis of oils and fats for linoleic and linolenic acids, and its extension by Beadle and Kraybill[2] to arachidonic acid, involves heating 0.1 gram of sample in 10 ml of 1.3*N* potassium hydroxide-ethylene glycol solution for 25 minutes at 180°C, cooling, diluting to 250 ml with 99 percent ethanol, settling, filtering, and measuring the ultraviolet absorption of the final solution against that of a reagent "blank" solution treated in the same way.

The principal difficulties encountered in attempting to apply this method to materials having very low proportions of non-conjugated polyunsaturated fatty acid constituents were: (*a*) the KOH-ethylene glycol isomerization medium develops a strong and variable ultraviolet absorption on heating in the presence of air; (*b*) the background absorption by extraneous compounds is relatively high (Fig. 2); and (*c*) the proportion of conjugated (particularly diene) to non-conjugated components is relatively high, raising the question as to what happens to the conjugated components during the isomerization treatment.

The first difficulty was usually manifested by poor reproducibility of analyses for linolenic and arachidonic acids. In some cases the observed spectral densities were actually negative, indicating that the density of the solute was less than the difference in density between two blanks run in the constant-temperature bath at the same time, thus making impossible the accurate analysis for small proportions of these constituents.

Improved transparency of the reagent solution was obtained by protecting the solutions during heating with a nitrogen atmosphere, but the reproducibility of absorption was not greatly improved in the first attempts. Further experiments along this line were abandoned when it was found that substitution of glycerol for ethylene glycol as a saponification and isomerization medium overcame previous difficulties without the added complication of excluding air.

The absorption and reproducibility of heated alkaline glycerol and ethylene glycol solutions are compared in Fig. 4 and Table II. The substantial gain in reproducibility and transparency by use

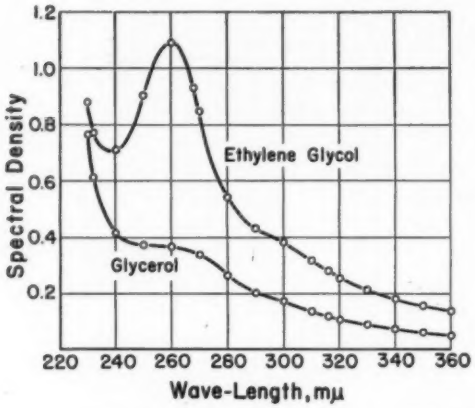


FIG. 4. Absorption curves for 1.30*N* KOH ethylene glycol and 11 percent KOH-glycerol solutions heated in the presence of air for 25 minutes at 180°C. Spectral densities are for 1-cm layers of 10 percent solutions in absolute ethanol, compared with a 1-cm layer of absolute ethanol.

of glycerol, together with application of corrections to isomerized samples for absorption by extraneous compounds, has made possible greatly improved sensitivity and accuracy in analysis for linolenic and arachidonic acids. The relatively high absorption below 230 mμ, caused by the potassium hydroxide in the solutions, is not troublesome since relatively high dilutions of the 10 percent solutions are required for absorption measurements on isomerized fatty materials in the diene region.

A modified analytical procedure applicable to a wide variety of oils, fats, soaps, and other fatty materials was devised. Standardization was accomplished by following spectrophotometrically the effect of time of heating on the isomerization reaction, using purified linoleic and linolenic acids.

The linoleic acid standard, prepared from corn oil by standard methods, had an iodine value of 181.9 and an acid value of 198.9. Spectrophotometric examination indicated the presence of approximately 0.3 percent linolenic acid, 1.7 percent conjugated diene, and 0.02 percent conjugated triene acids, and hence a purity not greater than 98.0 percent. The linolenic acid standard, prepared from perilla oil by standard methods, had an iodine number of 273.5 and an acid value of 200.6. Spectrophotometric examination indicated the presence of approximately 1.5 percent conjugated diene and 0.1 percent non-conjugated tetraene acids, and hence a purity not exceeding 98.4 percent. No purified arachidonic acid was available for the standardization tests in glycerol.

Curves showing relation of time of heating at 180°C to specific extinction coefficients (in absolute ethanol) at the absorption maxima of the conjugated products formed are shown in Fig. 5 for 0.1-gram portions of the pure acids each in 11.0 grams of glycerol solution containing 11.0 percent KOH by weight. Although a heating time of about 45 minutes would appear to be optimum for analytical purposes, a slightly less favorable time of 30 minutes was actually selected in order to use the data of Beadle and Kraybill[2] for arachidonic acid. The observed coefficients, their values adjusted to 100 percent purity of the acids, and for comparison the coefficients reported by Beadle and Kraybill,[2] are shown in Table III. Complete absorption curves for the treated acids are shown in Fig. 2.

In order to determine the effect of the alkali treatment on conjugated fatty constituents, measurements were made under these same conditions on conjugated materials, with the results shown in Fig. 6 and Table IV. It is apparent that the conjugated acids and glycerides are little affected by this treatment. Hence in the analysis of a sample for non-conjugated constituents by the alkali-isomerization method, correction must

TABLE II. Absorption and reproducibility for 1.30*N* KOH-ethylene glycol and 11 percent KOH-glycerol solutions heated in the presence of air for 25 minutes at 180°C. Spectral densities for 10 percent solutions in absolute ethanol, 1 cm at 260 mμ, 5 cm at 316 mμ (i.e., thicknesses and concentrations commonly used for analysis of animal fats) *vs.* equal layers of absolute ethanol. Each group of six samples was run simultaneously in the constant temperature bath.

	Spectral density			
	KOH-ethylene glycol		KOH-glycerol	
	260 mμ	316 mμ	260 mμ	316 mμ
	1.13	2.79	.359	.585
	1.18	2.91	.362	.590
	1.02	2.69	.377	.600
	1.09	2.71	.377	.605
	1.22	2.88	.377	.605
	0.96	2.51	.370	.595
Mean	1.10	2.75	.370	.597
Standard deviation	±0.10	±0.15	±.008	±.008

be made for conjugated constituents originally present in appreciable amounts.

In addition, correction of ultraviolet absorption data on the isomerized product must be made for background absorption by extraneous compounds (see Fig. 2 for absorption curves of typical alkali-isomerized fatty materials). This background absorption may be corrected for in the same way as proposed above in the analysis for conjugated constituents.

Thus after the isomerization treatment, absorption near 232 mμ will be contributed by conjugated diene components originally present in the fatty material; by conjugated diene components from isomerized linoleic, linolenic, and arachidonic acid compounds; and by extraneous compounds. Absorption near 268 mμ will be contributed by conjugated triene components originally present; by conjugated triene components from isomerized linolenic and arachidonic acid constituents; and by extraneous compounds. Absorption near 316 mμ will be contributed by conjugated tetraene components originally present; by conjugated tetraene components from isomerized arachidonic acid compounds; and by extraneous compounds.

In the equations which follow, all primed coefficients refer to the isomerized product in absolute ethanol, and unprimed coefficients refer to the untreated material dealt with in Eqs. (1)–(6); the subscripts have the same meaning as before.

The coefficient at 232 mμ corrected for conjugated diene material originally present is,

TABLE III. Specific extinction coefficients for alkali-isomerized pure polyunsaturated fatty acids.

Sample No.	Acid	Wave-length, mμ	Specific extinction coefficients [a]	[b]	[c]
7	Linoleic	232	89.1	88.9	86.0[d]
8	Linolenic	232	60.8	60.0	60.9[d]
8	Linolenic	268	52.6	53.4	53.2
—	Arachidonic	232	—	59.3	59.3[d]
—	Arachidonic	268	—	53.4	53.4
—	Arachidonic	316	—	22.6	22.6

[a] Observed values, in glycerol 30 minutes at 180°C.
[b] Values used for analysis, in glycerol 30 minutes at 180°C: observed values adjusted to 100 percent purity of acid; arachidonic acid coefficients assumed valid for these conditions.
[c] Values reported by Beadle and Kraybill (see reference 2) in ethylene glycol 25 minutes at 180°C.
[d] At 234 mμ.

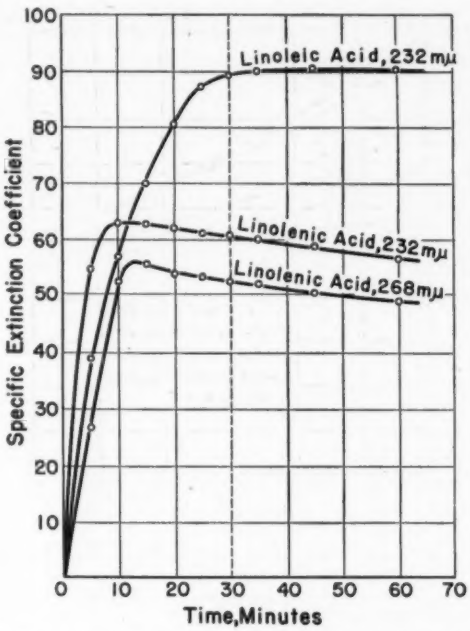


FIG. 5. Specific extinction coefficients in absolute ethanol for 0.1-gram samples of purified linoleic and linolenic acids heated for various lengths of time at 180°C in 11 percent KOH-glycerol solution.

approximately:

$$k'_2 = k'_{232} - k_{232} + 0.04 \text{ for fats and esters,} \quad (7)$$

$$k'_2 = k'_{232} - k_{232} \text{ for soaps and acids,} \quad (7a)$$

where k'_{232} and k_{232} are the observed specific extinction coefficients of the material after and before isomerization. The constant 0.04 compensates for the fact that fats and esters are converted to soaps by the hot alkali treatment (see Eqs. (1) and (1a)). This term is important only for samples containing less than about 1 percent of linoleic acid.

The coefficient at 268 mμ corrected for background absorption is approximately:

$$k'_3 = 4.1[k'_{268} - \tfrac{1}{2}(k'_{262} + k'_{274})], \quad (8)$$

where k'_{262}, k'_{268}, and k'_{274} are the observed specific extinction coefficients, and the constant 4.1 was evaluated from measurements on ethanol solutions of the isomerized pure linolenic acid preparation 8. No explanation is advanced at this time for the fact that the value of this constant is

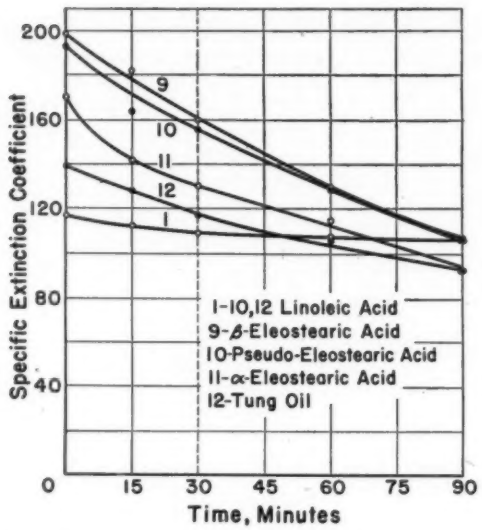


FIG. 6. Specific extinction coefficients in absolute ethanol for conjugated fatty acids and tung oil before and after heating for various lengths of time at 180°C in KOH-glycerol solution.

considerably higher than was obtained for β-eleostearic acid used in Eq. (2).

The coefficient at 268 mμ corrected for conjugated triene material originally present is approximately:

$$k''_3 = k'_3 - k_3, \qquad (9)$$

where k'_3 is given by Eq. (8) and k_3 by Eq. (2). This correction is usually small, and a coefficient of unity for k_3 may be used instead of the value 0.8 indicated by Table IV to compensate for the application in Eq. (8) of the factor 4.1 to the undestroyed conjugated material as well as to that formed by isomerization.

The coefficient at 316 mμ corrected for background absorption is approximately:

$$k'_4 = 2.5[k'_{316} - \tfrac{1}{2}(k'_{310} + k'_{322})], \qquad (10)$$

where k'_{310}, k'_{316}, and k'_{322} are the observed specific extinction coefficients and the constant 2.5 was evaluated from isomerized arachidonic acid.[2,13] The final corrected coefficient at 316 mμ, is assumed to be:

$$k''_4 = k'_4 - k_4, \qquad (11)$$

where k'_4 is given by Eq. (10) and k_4 by Eq. (3).

Since the specific extinction coefficient of a mixture at a selected wave-length is equal to the sum of the specific extinction coefficients of the components, each multiplied by its proportion in the mixture, simultaneous equations may be set up for the system and solved for the proportion of each component. Using the standardized data of Table III and the corrected specific extinction coefficients defined by Eqs. (7) or (7a), (9), and (11), these final equations for the KOH-glycerol procedure, with 30 minutes heating at 180°C, are:

$$x = 1.125k'_2 - 1.27k''_3 + 0.04k''_4, \qquad (12)$$

$$y = 1.87k''_3 - 4.43k''_4, \qquad (13)$$

$$z = 4.43k''_4, \qquad (14)$$

where x, y, and z are the concentrations, expressed as percent fatty acid, of linoleic, linolenic, and arachidonic acid constituents present in a given sample. If corrections are omitted for absorption by undestroyed conjugated and other extraneous constituents, the coefficients k'_2, k''_3, and k''_4 become the observed coefficients k'_{232}, k'_{268}, and k'_{316}, and the calculations indicated by the final equations become equivalent to those described by Mitchell *et al.*,[1] and Beadle and Kraybill.[2]

PROCEDURE FOR ANALYSIS

A sample of oil, melted fat, acid, ester, or powdered soap about 200 mg in size is weighed to the nearest milligram, dissolved in about 75 ml of neohexane[14] or in 95 percent ethyl alcohol,[15]

TABLE IV. Specific extinction coefficients in ethanol, and wave-lengths of absorption maxima for conjugated fatty materials before (k) and after (k') heating for 30 minutes at 180°C in KOH-glycerol solution. The ratio k'/k gives approximately the proportion of conjugated material undestroyed by the treatment.

Sample No.	Material	Before		After		
		λ	k	λ'	k'	k'/k
1	10,12 linoleic acid	231	117	231	110	.94
9	β-eleostearic acid	268	200	268	161	.81
10	Pseudo-eleostearic acid	267.5	193	267.5	156	.81
11	α-eleostearic acid	270	171	269	131	.77
12	Tung oil	270	140	269	118	.84

[14] Neohexane, Phillips Petroleum Company, technical grade. Isooctane (2,2,4-trimethylpentane), Rohm and Haas, Bureau of Standards certified grade, is also a satisfactory solvent. The transparency of these solvents was improved greatly by passage through a column of silica gel; see M. M. Graff, R. T. O'Connor, and E. L. Skau, Ind.

TABLE V. Analysis of tallow 15 for polyunsaturated fatty acids by different methods. [(*a*), (*b*), (*c*) were isomerized in KOH-glycerol; (*d*) was isomerized in KOH-ethylene glycol (no corrections made).]

Acid	(*a*) Proposed method, percent	(*b*) Correction for conjugation neglected, percent	(*c*) Corrections for background and conjugation neglected, percent	(*d*) Method of Mitchell *et al*[a] Beadle *et al*,[b] percent	(*e*) Method of Bradley and Richardson,[c] percent	(*f*) Method of Brode *et al*,[d] percent	(*g*) Method of Kass,[e] percent
Linoleic	3.39 3.39	4.55 4.53	4.25 4.23	4.35 4.15	—	—	—
Linolenic	.61 .64	.63 .66	.74 .75	.51 .57	—	—	—
Arachidonic	.68 .64	.70 .67	1.04 1.02	1.20 1.14	—	—	—
Conjugated diene	.86	—	—	—	.90	.90	.94
Conjugated triene	.011	—	—	—	.12	.14	.094
Conjugated tetraene	.002	—	—	—	.030	.030	.040

[a] See reference 1.
[b] See reference 2.
[c] See reference 3.
[d] See reference 7.
[e] See reference 8.

transferred quantitatively to a 100-ml volumetric flask, and made up to volume. Spectral densities are then measured on a Beckman spectrophotometer preferably at 2-mμ intervals from 324 to 306 mμ, 282 to 260 mμ, and 240 to 226 mμ, dilutions and cell lengths (1 to 5 cm) being adjusted so that, whenever possible, observed densities are between 0.2 and 0.8. Curves are plotted for examination, the necessary specific extinction coefficients are calculated, and the percentages of conjugated diene, triene, and tetraene fatty acid components present are computed by Eqs. (1) through (6).

Two additional samples of about 100 mg are weighed to the nearest half milligram in Pyrex-glass cups of approximately 1-ml capacity; these samples are for analysis in duplicate for nonconjugated constituents. The isomerization reagent is prepared by adding 17.5 grams of KOH (A.C.S. standard) to 100 ml of glycerol (U.S.P. XII) and heating to 200°C with constant stirring to dissolve the alkali and drive off excess water. The resulting reagent should have a KOH concentration of 10.9 to 11.0 percent by weight, and may be checked by titration. Three 6-×1-inch Pyrex-glass test tubes, serving as reaction tubes, are filled with 11.0-gram portions of the solution and then suspended to a depth of 4½ inches in a constant temperature bath operated at 180°C±0.5°.

When the temperature of the solution in a tube has reached 180°, the glass vessel containing a weighed sample is dropped into the tube, the tube is removed from the bath, swirled vigorously for one minute, then returned to the bath and covered with a crucible cover. At the end of a minute's heating in the bath, the tube is removed and inspected.[16] If the solution is clear, the tube is returned to the bath. If saponification or solution is not complete, the tube is again swirled a few times, the cover is replaced, and the tube returned to the bath. The procedure is repeated with the duplicate sample. An empty sample container is dropped in the third reaction tube, which serves as a reference blank.

Eng. Chem. Anal. Ed. **16**, 556 (1944); also Nat. Bur. Stand. Tech. News Bull. No. 325, p. 37 (May 1944).

[15] Synthetic absolute methyl alcohol is also a suitable solvent. We are indebted to E. L. Borg of the United States Rubber Company for suggesting its use.

[16] It is essential that all samples saponify or dissolve within one to two minutes after being dropped into the reaction tube. Animal fats and fatty acids saponify within a minute. Monohydric esters saponify less readily. This may be remedied by adding 0.1 percent of pure palmitic or stearic acids to the KOH-glycerol reagent after the preliminary heating at 200°. In the case of soap samples, prompt solution is accomplished by wetting the sample with 3 drops of water or glycerol, stirring it to a homogeneous mass, and dropping sample and glass stirring rod into the reaction tube.

TABLE VI. Analysis of miscellaneous fatty materials for polyunsaturated constituents. Effect of variations in isomerization medium, size of sample, and temperature; reproducibility of results; and analysis of known mixtures.

Sample	No.	Iodine value	Non-conjugated acids, percent			Conjugated acids, percent			Total polyunsaturated acids, percent	
			Linoleic	Linolenic	Arachidonic	Diene	Triene	Tetraene	Spectrophotometric	Chemical
Tallow	13	57.5	4.8	0.46	0.29	0.68	0.008	0.001	6.24	8.3[a]
Tallow, hydrogenated	14	49.4	0.06	0.00	0.00	0.15	0.000	0.000	0.21	—
Oleic acid, purified	20	91.1	1.05	0.04	0.00	0.52	0.000	0.000	1.61	—
Perilla oil	18	207.3								79.9[b]
isomerized in glycerol			13.1	55.6	1.3	2.2	0.07	0.02	72.3	
isomerized in ethylene glycol			13.1	55.1	1.1	—	—	—	71.6	
Soybean oil, coldpressed	19	—								—
mean of 20 runs			43.25	4.12	0.98	2.6	0.01	0.000	51.0	
standard deviation			±0.21	±0.09	±0.02					
Tallow	17	56.7								8.8[a]
mean of 20 runs			4.66	0.52	0.23	0.90	0.02	0.000	6.33	
standard deviation			±0.09	±0.03	±0.03					
Tallow	21	—								—
0.1 g[c]			4.1	0.48	0.19	—	—	—		
0.2 g			3.8	0.46	0.18	—	—	—		
0.3 g			3.0	0.44	0.18	—	—	—		
Linoleic acid concentrate	22	168.0								—
175°C			62.2	6.8	0.26	—	—	—		
180°[c]			64.5	7.1	0.25					
185°			66.5	7.2	0.21					
Mixture	23	—								—
Calculated			9.1	0.98	0.00	0.36	0.001	0.000	10.44	
Found			9.1	0.90	0.00	0.32	0.001	0.000	10.32	
Mixture	24	—								—
Calculated			4.6	0.50	0.00	0.18	0.0004	0.000	5.28	
Found			4.9	0.40	0.00	0.18	0.0005	0.000	5.48	
Mixture	25	—								—
Calculated			0.48	0.05	0.00	0.02	0.000	0.000	0.55	
Found			0.58	0.05	0.00	0.02	0.000	0.000	0.65	

[a] Thiocyanometric method, assuming the presence of saturated, oleic, and linoleic acids only.
[b] Thiocyanometric method: linoleic acid, 14.7 percent; linolenic acid, 65.2 percent.
[c] Standard.

Thirty minutes after the sample is dropped into the reagent tube, the tube is removed and cooled under tap water, 20 ml of absolute ethanol[15] is added, and the mixture is stirred. The tube is placed in a beaker of hot water over low heat, and stirring is continued until the viscous product dissolves completely. The solution is then transferred quantitatively to a 100-ml volumetric flask and made up to volume with absolute ethanol.[15] No settling and filtering are required when absolute alcohol is used as a solvent. As little as 1 percent water in the solvent causes a precipitate.

Spectral density measurements are made exactly as specified in the analysis for conjugated constituents, except that the "blank" is the heated reagent and must be diluted to the same extent as the sample solutions. The required specific extinction coefficients are calculated, and the percentages of linoleic, linolenic, and arachidonic acids in the sample are computed by means of Eqs. (7) through (14).

If the calculated value for k'_3 in Eq. (8) is zero or negative, no linolenic acid is present. A further criterion is used when small positive values for k'_3 occur: if the plotted curve of density *vs.* wavelength shows no evidence of simultaneous absorption maxima or inflections near wave-lengths 268 and 278 mμ, linolenic acid is absent.

The complete analysis of a tallow sample for

polyunsaturated fatty acid constituents is shown in Table V. The data show (*a*) results by the proposed method; (*b*) the effect on the results of omitting correction for undestroyed conjugated constituents; (*c*) the effect of neglecting in addition corrections for background absorption; (*d*) results obtained for the non-conjugated acids by previous methods;[1,2] and (*a*), (*e*), (*f*), (*g*) comparison of different methods for the analysis of conjugated constituents.

The data illustrate that, in fats of this type (1) the correction is appreciable for the presence of undestroyed conjugated diene acids, but negligible for conjugated triene and tetraene acids; (2) large errors may be made in the results if background absorption is neglected; and (3) the methods and standardization data of other investigators for conjugated acid assay give essentially the same results but are probably greatly in error, at least for triene and tetraene acids. The dependent nature of the simultaneous analysis for linoleic, linolenic, and arachidonic acids should be kept in mind; for example, any large error in the determination of arachidonic acid (the only independently determined non-conjugated acid) will produce an appreciable error in the values calculated for linolenic and linoleic acids.

Results on the spectrophotometric analysis of some typical fatty materials for polyunsaturated constituents are shown in Table VI. Iodine values and percent linoleic acid obtained by the thiocyanometric method are shown where data were available for the chemical method. The spectrophotometric analyses are reported as percent acid in total sample, whether the sample be a fat, ester, or acid. The proportions of oleic and saturated acids can be calculated from the proportions of polyunsaturated constituents indicated (corrected to the basis of percent acid in total fatty acids), their theoretical iodine numbers, and the iodine number of the sample.

Mild selective hydrogenation (sample 14) reduces the linoleic and linolenic acid content of tallow to extremely low levels, and completely removes tetraene and triene conjugated constituents. The data for purified oleic acid (sample 20) illustrate the value of the method in detecting low amounts of polyunsaturated impurities in fatty materials used in research studies, for example, on oxidation. Neglect of background absorption in cases such as 14 and 20 would lead to completely erroneous results.

Application of the method to vegetable oils having substantial proportions of linoleic and linolenic acid components is also shown (samples 18 and 19). The principal value of the modified method in such cases is to furnish more reliable data for the less abundant constituents. The presence of tetraenoic fatty acid constituents in these oils is of interest.

In the case of the perilla oil sample, results of analyses in which both glycerol and ethylene glycol were used as isomerization solvents are in good agreement. The value for arachidonic acid obtained by the glycerol procedure is more reliable for reasons previously discussed.

Data in Table VI (sample 21) indicate that the size of the sample used in the analysis for non-conjugated constituents is rather critical, and should not differ from the standardized 0.1-gram size by more than about 10 percent. Similar deviations have been obtained with samples of soybean oil and partially hydrogenated tallow, indicating the need for further study of the isomerization reaction.

Tests on a tallow and a linoleic acid concentrate (sample 22) indicate that the temperature at which the isomerization reaction is carried out should be maintained constant to about $\pm 0.5°C$.

The high reproducibility of the method is illustrated by the standard deviation of 20 analyses (runs in duplicate on ten different days) for samples of soybean oil and tallow. All other results shown in Table VI are the average of duplicate analyses. No difficulty has been experienced in attaining this reproducibility when the procedure specified is closely followed and when saponification or solution occurs within one or two minutes after the sample is dropped into the reaction tube.

The accuracy of the method was tested by analyzing purified methyl esters of linoleic, linolenic, and oleic acids, and their mixtures. The calculated and found values of the diene and triene constituents are shown in Table VI for three mixtures which roughly simulate the composition of lard, tallow, and selectively hydrogenated tallow. The agreement between calculated and observed values is good.

The accuracy of the method is probably not so high as the reproducibility. Factors contributing to the uncertainty of results include imperfections in correction for interfering constituents, incomplete identification of impurities in the standards used, and the possible presence in some samples of isomers having conjugation rates different from those of the standards. Agreement with the chemical method is best when linoleic acid is present in much greater abundance than acids of higher unsaturation (e.g., lard and tobacco seed oil). It is conservatively estimated that the errors of the method are within ±10 percent of the quantity present when that quantity is near 10 percent, ±25 percent near 1 percent, and that the results are at least correct in order of magnitude when the quantity present is 0.1 percent or less.

The application of the method to a wider variety of fatty materials will be published elsewhere.[17]

SUMMARY

A method is described for the simultaneous spectrophotometric determination of non-conjugated and conjugated diene, triene, and tetraene fatty acid constituents in vegetable oils, animal fats, their soaps, and purified fatty acid preparations. The method involves measurement of the ultraviolet absorption of a sample before and after isomerization. Correction of the data for absorption by extraneous compounds and use of alkaline glycerol as an isomerization medium result in greater sensitivity and accuracy in the analysis for minor and trace proportions of these polyunsaturated constituents.

[17] B. A. Brice, M. L. Swain, B. B. Schaeffer, and W. C. Ault, *Oil and Soap* (in press).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 8 AUGUST, 1945

An Investigation on the Formation of Surface Layers*

D. D'EUSTACHIO AND S. B. BRODY
Bliley Electric Company, Erie, Pennsylvania

(Received June 14, 1945)

When a series of back reflection x-ray powder pictures of a freshly crushed coarse crystalline powder is taken, successive pictures exhibit a progressive change. Individual spots sharpen, diffuse portions of powder rings disappear, and the number of spots increases markedly with time. The increase in spots is not caused by breaking up of the particles. An account of the technique used for observing the effect is given. Experiments are described which indicate that the changes are definitely associated with the surface. The time rate of change increases with the hardness of the material and appears to be related to the degree of valance binding in the structure. A tentative explanation of the phenomenon is proposed which can account for all the observed effects.

INTRODUCTION

SEVERAL workers[1,2] have reported that the x-ray reflections from single crystals decrease in intensity and become sharper upon etching the surface. This has been interpreted to mean that the surface layer is poorer crystal than the interior. From the tripling of Laue spots in transmission photographs of ground quartz slabs,

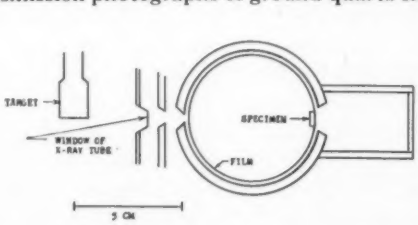


FIG. 1. A schematic diagram showing general arrangement of apparatus.

* A preliminary report of this work was presented at the Pittsburgh meeting of the American Physical Society in April, 1944.

[1] M. Y. Colby and Sidon Harris, Phys. Rev. **43**, 562 (1933).

[2] R. M. Bozorth and F. E. Haworth, Phys. Rev. **45**, 821 (1934).

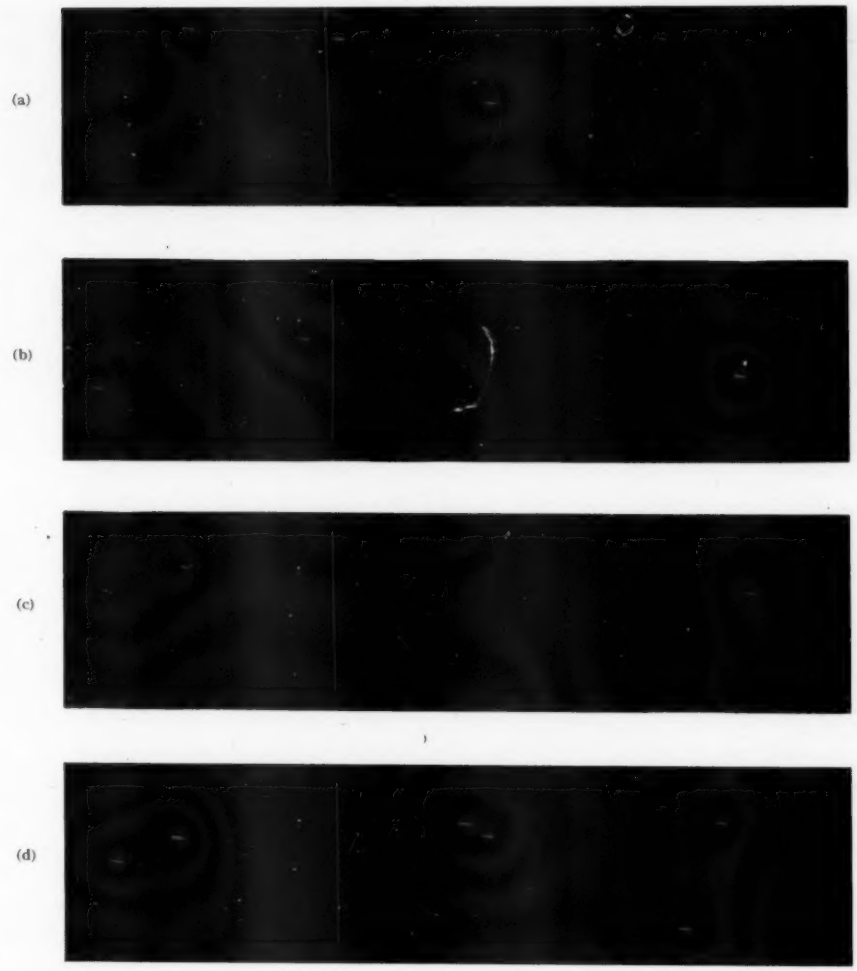


FIG. 2. X-ray powder photograph of calcite (a) 5 minutes, (b) 40 minutes, (c) 2 hours, (d) 22 hours after crushing.

Murdock[3] concluded that the interior was a single crystal and the surfaces "mosaic." Armstrong[4] reports surface material with disorientations up to 3° in quartz. It has been generally believed that this disturbed surface is produced by the mechanical forces involved in the grinding and polishing operations. It is difficult to understand how such large disorientations can be produced by mechanical forces acting on the crystal surface during lapping, since the grinding of a relatively brittle crystal like quartz consists of a succession of small fractures. This paper reports experiments which indicate that these "disoriented" surfaces may be formed at some time subsequent to fracture of the crystal.

[3] C. C. Murdock, Phys. Rev. **45**, 117 (1934).
[4] E. H. Armstrong, paper presented at the Cleveland meeting of the A.A.A.S., September, 1944.

EXPERIMENTAL METHOD

In order to keep the exposure times as short as possible, it was necessary to use an x-ray camera having the largest possible angular aperture and minimum target-to-film distance.

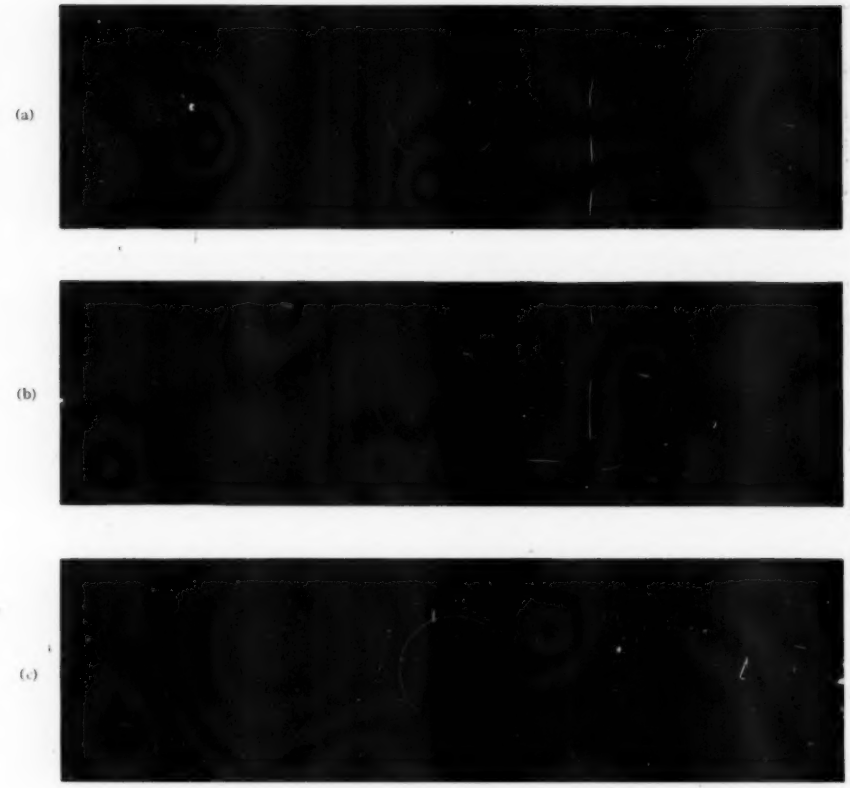


FIG. 3. X-ray powder photograph of topaz (a) 4 minutes, (b) 15 minutes, (c) 60 minutes after crushing.

At the same time it was desirable to have adequate dispersion. It was found that these conditions were sufficiently well met with a back reflection focusing camera $2\frac{1}{2}''$ in diameter. Figure 1 shows the experimental arrangement schematically. With this arrangement an area of crystalline powder 3 mm wide by 8 mm high was irradiated by the x-ray beam. In the earlier experiments, a water cooled G.E. copper tube was used, operating at approximately 40 kv and 7 ma. Later work was done with an air cooled Phillips copper tube operating at 38 kv and with currents between 6 ma and 8 ma. Exposures were 3 to 5 minutes. The change that occurs in this length of time is small when it is compared with the total change that takes place.

Specimens were prepared in the following manner. A small quantity of the material to be investigated was crushed in a mortar. Particles between 200 and 325 mesh were selected by successive screening. In the earlier experiments a dry screening technique was used. To secure more accurate particle sizing wet screening was used in the later work, and the limits of particle size obtained by this technique were checked by examining samples with an optical microscope. The number of particles finer than 350 mesh was negligible. A layer of sifted powder 5 mm by 10 mm by 1 mm was spread on a thin card with a mechanical fixture that insured uniform thickness of the material and cemented in place with a $\frac{1}{2}$ percent solution of ethyl cellulose in ether. The mounted specimen was held in position in the camera by a clip and two register pins arranged to facilitate rapid loading and at the

(a) (b)

FIG. 4. Enlargement of a series for a quartz specimen (a) 3 minutes, (b) 40 minutes, after crushing.

same time guarantee that the card did not shift while films were changed. The time which elapsed from the start of the crushing operation to the beginning of the first exposure was always less than ten minutes and often as little as two minutes. After the first picture a number of other pictures were taken at suitable intervals. Special precautions were taken not to disturb the specimen surface between exposures. An accurate register device on the x-ray unit was provided for returning the camera to the same position. Films in any one series were developed together after the last exposure had been made.

RESULTS

Typical series are reproduced in Figs. 2 and 3. Figure 2 shows calcite and Fig. 3 shows topaz. The ages given in the legends are the total elapsed times from the beginning of the crushing operation to the start of the exposure. No significant error is involved in counting the age from the beginning of the crushing operation since the time consumed in this process was generally less than half a minute.

TABLE I.

Substance	Relative rate of change
Topaz	1
Quartz	2
Calcite	2
Barite	2.5
Anhydrite	2.5
Adularia	3

The most significant feature of the calcite and topaz series is the marked decrease, with time, in the diffuseness of the rings. What appear to be imperfect, broken powder rings with only a few spots, in the earliest pictures, change into strings of discrete well defined spots with a marked decrease in continuous background. In some of the later pictures the background of diffuse portions of rings has disappeared completely.

In addition to the diffuse rings there appear in the first pictures of each series numbers of spots which also appear in the subsequent pictures of the same sample and can be superimposed. This is a feature of all series taken and was useful in making sure that the powder specimen had not shifted or been disturbed in any other way.

A view of a typical detail of the change which is occurring is given in Fig. 4. Two enlargements of the same small region, taken from a series of photographs of quartz, are shown. It will be noted that the sharpness of the spots increases and the intensity declines in the later picture. At the same time diffuse patches become resolved into spots.

The experiments were repeated with several materials, and gave similar results. In each instance the changes were more rapid at first. For all the material except rocksalt the changes finally became insignificantly small after enough time had elapsed. The behavior of rocksalt has not been studied sufficiently; after 42 hours

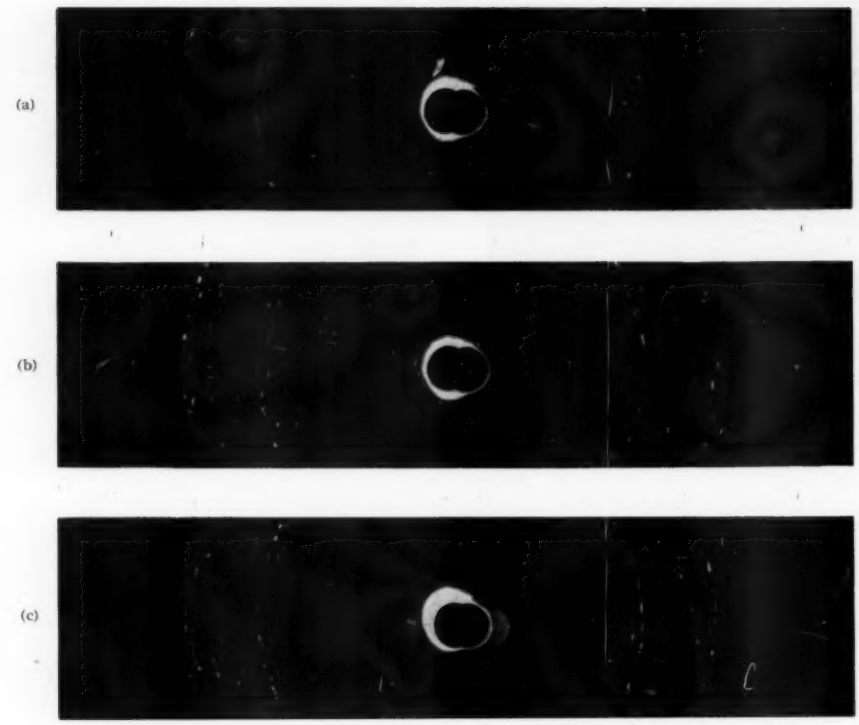


FIG. 5. X-ray powder photograph made with rescreened material (a) 4 minutes, (b) 22 hours after crushing, (c) portion of material rescreened 22 hours after crushing and photographed 4 minutes after rescreening.

small changes were still barely perceptible and the experiment was discontinued. Experiments were repeated numerous times and the rate of change is definitely a characteristic of the compound. Table I gives the relative rates of change for a number of substances.

To aid in keeping the exposure time short, most of the work was done with unfiltered radiation. The pictures were compared with standard powder patterns obtained from finely powdered material, with a nickel filter and eight hours' exposure. All but a few isolated spots in the pictures of coarse powders taken with unfiltered radiation fell on the powder rings of the standard pictures. With the short exposures used, Laue reflections do not appear often, and it is not very likely that many Laue spots would fall fortuitously on characteristic radiation rings. When pictures of the same "old" specimen were taken with and without a nickel filter 0.00035" thick, the only noticeable effect of removing the filter, besides increases in intensity, was that some of the stronger spots became doubled (K_α and K_β).

Making the assumption that all the spots on a given ring are reflections of the characteristic radiation, one may readily calculate the probable number of spots per ring that should appear if all the particles in the irradiated volume of the sample were homogeneous single crystals. In the case of topaz, for instance, allowing for absorption, the number of particles of the specimen which could contribute to the diffraction was around 3000 to 5000 under the experimental conditions described. The probability of any single crystal contributing to a given ring is given by[5] $j/2 \cos\theta d\alpha$, where j is the multiplicity, θ is

[5] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Company, Inc., New York, 1935), p. 417.

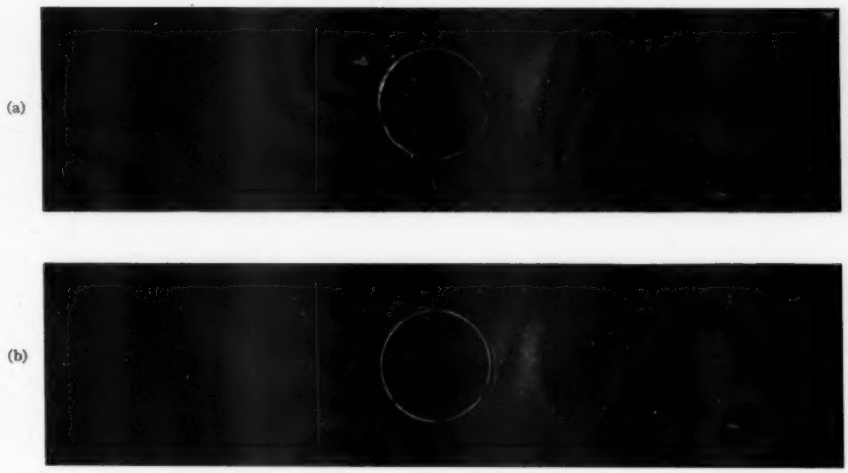


FIG. 6. X-ray powder photograph of barite ($BaSO_4$) (a) 8 minutes, (b) 2 hours after crushing.

the Bragg angle, and $d\alpha$ is the small angle through which the crystal continues to reflect. Substituting the appropriate values for the inner rings in the topaz series one obtains approximately 1/1000 for the probability of a particle of arbitrary orientation contributing to one of these rings. There should therefore be only a few spots per ring, if the particles are single crystals. In the early pictures (see Fig. 2a, 3a), there are generally discernible only five or six well defined spots per ring. But the number of spots increases greatly in subsequent pictures and the final pictures of some series have hundreds of spots per ring.

That this huge increase in the number of spots is not due to the particles breaking up was determined in another experiment. A freshly crushed and sifted sample was divided into two parts. The first part was mounted immediately and the usual series of photographs was taken. The second portion was saved. When the pictures of the first portion showed that no further changes were taking place, the second part was resifted thoroughly in the 325 mesh screen. As is inevitable in all resiftings even when the original sifting has been carried out with care, a small amount of material fell through the screen. No fine powder was found on sifting. The small amount of material consisted of particles hardly smaller than 325 mesh.

In all experiments with material that was re-screened after standing, the range of particle size used was 200 to 325 mesh, a ratio of less than two to one. Hence any material which had broken off from the particles should have fallen through the finer screen. Some of the rescreened material was mounted, and a series of pictures taken in the usual way, over a period of hours. These pictures were all indistinguishable from each other and very like the final photograph of the first portion.

Pictures taken with the residue, after the second screening had just about as many spots, on the average, as pictures taken with the first portion of the sample. Hence, the very large increase in the number of spots which occurs in all the later pictures cannot be caused by the particles breaking up. Figure 5 shows the first and last pictures of the first half of the sample compared with a picture of the rescreened portion. Several series of pictures were taken, with samples that had been stored for various lengths of time (up to 10 days) after sifting. No changes were ever observed in old powders.

A series of specimens was made with successively smaller particle sizes for each succeeding specimen and the exposure time needed to obtain the same degree of blackening in the more intense spots was investigated. If one assumes that the spots which occur on the earliest photo-

FIG. 7. X-ray powder photograph of anhydrite ($CaSO_4$) (a) 5 minutes, (b) 4 hours after crushing.

graph of a series and recur in the same positions in the later films are due to reflections from the body of the crystal, and that the spots which only appear on later pictures and become sharper with time are associated in some manner with the surface of the particles, one would expect the spots of the first kind to decrease in intensity with decreasing particle size, but the intensity of spots of the second kind to be independent of particle size. It was found that for spots of the first kind, the exposure times decreased with increasing particle size. The exposure time required for equal blackening of the second type of spot was independent of specimen particle size, but the *number* of this type of spot recorded on the film decreased as the diameter of the particles was increased.

Another experiment was carried out, to separate more conclusively the effects associated with the surface from those connected with the interior of the particles. The original experiment was repeated, first with anhydrite ($CaSO_4$) and then barite ($BaSO_4$). One would expect the more highly absorbing substance to show a relatively greater contribution from the surfaces of the particles. The mass absorption coefficient for copper $K\alpha$ radiation in barite is about 1000 cm^{-1} compared with 230 cm^{-1} for anhydrite. A path length of approximately 10^{-3} inch in barite is sufficient to reduce the intensity to 10 percent of its original value. The corresponding path length in anhydrite is of course approximately four and one-half times as great. Since the particles in the sample were 200 to 325 mesh, a surface layer of each barite particle of depth roughly a tenth to a fifth of the radius contributed to the reflections. On the other hand, the whole interior of the anhydrite particles was available for diffraction. The pictures of anhydrite and barite are shown in Figs. 6 and 7. Comparing them we see that almost no individual sharp reflections occur in the first barite picture, while the second barite picture contains many. The first anhydrite picture shows approximately the proper number of individual reflections to be expected from this material and the second anhydrite picture shows the usual increased number of well defined spots. The number of new sharp spots is about the same in Figs. 6 and 7. This is exactly the effect to be expected from the high absorption of barite if the new spots on the later films are caused by the reflections of the x-rays from material on or near the surface of the particles.

CONCLUSION

The final interpretation of the work described above must await the results of further experiments. It is nevertheless useful to examine the physical meaning of the results obtained so far

and hazard a guess as to the mechanism that may be producing the changes. Since very little material can be separated on rescreening samples that have been stored for some time, the earlier suggestion[6] of spontaneous breaking up of the particles cannot be accepted as an explanation of the large number of spots occurring in photographs of old material. Repeated washing and even, if the particles are small enough, long etching does not change the character of the photographs obtained from old material. The conclusion that disoriented material is securely bound to the particles is inescapable. A satisfactory theory of the phenomenon must provide some reasonable mechanism for changing single-crystalline material into polycrystalline material. It should be noted however, that the randomness of orientation required of the polycrystalline material is not very large.

An acceptable theory must also account for the change from diffuse spots to sharp spots and the corresponding decrease in intensity that accompanies this change. Changes of this character require an increase in resolving power and can only be accounted for by an increase either in the number of "grating" spaces or the perfection of the "grating." The effect is not a disruption of the regular crystal structure such as assumed by Allison[7] to account for the increase in reflection coefficient that was observed when freshly cleaved calcite surfaces were polished.

The following mechanism, which is at least sufficient to account for all the observations noted thus far, is tentatively suggested. Whenever fracture occurs, part of the energy used to break the crystal remains on the surface in the form of structure irregularity. Not more than a very few surface layers are disturbed by the rupture. There may thus result a very thin highly stressed (possibly glassy) surface layer. This state explains the diffuseness of the pattern first observed. If one supposes that such a state is not stable, then atoms which most nearly occupy positions on the lattice of the mother crystal have less tendency to migrate than those which are badly displaced. As a result certain atoms or groups of atoms will act as nuclei for growth of surface "crystallites." The exact orientation of each little block of crystal will be a compromise between the position of the "nucleus" and the mother crystal. If this view is correct, the degree of sharpness finally attained by some of the initially diffuse spots indicates that the final linear dimension of the crystallites may be in the neighborhood of 50 to 100 times the unit cell translation.

It is hoped that single-crystal work now in progress will shed more light on the nature of the processes involved.

The authors wish to thank Dr. I. Fankuchen and the Chemistry Department of Polytechnic Institute of Brooklyn for the generous use of x-ray diffraction apparatus and laboratory facilities and Mr. Samuel Gordon of the Philadelphia Academy of Natural Sciences who very kindly supplied us with many excellent specimens of the minerals.

[6] Phys. Rev. **65** (1944), Abstract No. **B6**.
[7] S. K. Allison, Phys. Rev. **41**, 13, 688 (1932).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 35, NUMBER 8 AUGUST, 1945

New Etching Pattern of Quartz and Its Uses for the Determination of Electric Axes and the Detection of Crystalline Defects

CHOONG SHIN-PIAW
Institute of Radium, National Academy of Peiping, Kunming, China
(Received October 30, 1944)

A new etching pattern of quartz is formed, when a plate cut normal to the optic axis is subjected to the actions of hydrofluoric acid and electric field simultaneously. The general aspects of the pattern vary with the field intensity. In the case of weak field, the etching pattern consists of a number of regularly orientated small triangular figures which change their orientation from one twinned part to another. On the other hand, under the action of strong field, the triangular figures produced are generally smaller in size but by far greater in number than in the first case. The most striking features, exhibited in most of the etched surfaces under strong fields, are the appearance of groups of parallel long lines and bands as well as a kind of new etched zones which are similar in appearance but unlike in nature, the ordinary twinned zones. These groups of parallel lines and bands are exactly parallel to the electric axes of quartz; the presence of these lines and bands affords a very convenient way for the determination of the axes. There is reason to believe that the localities where the groups of parallel lines and bands as well as the new etched zones are found, are the seats of new crystalline defects arising from the so-called 'primary' crystal flaws. Some effects of the new crystalline defects on the oscillation characteristics of quartz oscillators are observed. The triangular etching figure, as revealed by a microscope, is composed of three sets of brushy parallel lines which make 120° with each other. The orientation of these sets of lines is a function of the applied field; those produced under weak field are approximately parallel to the three electric axes of quartz, with their outward ends directed to the positive extremities of the electric axes by compression, while those formed in strong field may deviate from the corresponding axes by as much as some ten degrees. Both the sense of deviation of the sets of lines with reference to the corresponding electric axes and the mode of arrangement of these brushy lines are found to be closely related with the sense of optical rotation of quartz. A general and accurate (to 1°) method for the determination of electric axes of quartz and their polarities, basing on the use of the orientations of the elementary etching figures, has been described.

INTRODUCTION

THE piezoelectric effect of quartz was discovered by J. and P. Curie[1] in 1880. Few applications were made of the effect, except in various instruments for laboratory work, until the piezoelectric properties of quartz were used by Langevin[2] in 1918 for detecting submarines and submarine mines, by Cady[3] in 1922 for the production of frequency standards, for frequency stabilization and as a means of generating electric oscillations at very constant radio frequencies. Since then, the use of piezoelectric quartz for wireless purposes has been developed very extensively.

Owing to the anisotropic nature of quartz, its use for making instruments often requires a precise knowledge about the homogeneity of the crystal as well as the directions of the optic and electric axes. Since the introduction of the method of etching for the detection of optical and electrical twinnings, great improvements in the test of homogeneity of quartz crystals have been realized. The optic axis of quartz may be easily determined by means of Norrenberg's polariscope or more accurately by the apparatus and procedure indicated by A. Arnulf.[4] The method commonly used for the determination of electric axes is to infer them from the characteristic crystalline form. This method, though simple, is not applicable to the case where the useful crystalline contour has been lost. A. de Gramont[5] and Pan Tcheng-Kao[6] had suggested to use the figures of corrosion, produced on the surface of a quartz plate cut perpendicular to the optic axis, for the determination of electric axes. According to these authors, the three heights of the triangular base of a projecting pyramid of corrosion are respectively parallel to the three electric axes of quartz and the vertexes of the base are directed towards the positive extremities of these axes by

[1] J. and P. Curie, Comptes rendus **91**, 294 (1880).
[2] P. Langevin, J. de Phys. **4**, 537 (1923).
[3] W. G. Cady, Proc. I. R. E. **10**, 83 (1922).
[4] A. Arnulf, Revue d'optique **10**, 453 (1931).
[5] A. de Gramont, Revue d'optique **10**, 213 (1931); *Recherches Sur le Quartz Piezoelectrique* (1935), p. 26.
[6] Pan Tcheng-Kao, Revue d'optique **10**, 153 (1931).

compression. Careful measurements carried out by the present writer have shown that the three basal heights of a pyramid of corrosion are directed along the mechanical axes rather than the electric axes. In a note[7] to *Nature*, the writer has indicated how the directions and the polarities of the electric axes can be deduced from such pyramids. But owing to the fluctuations in the orientation of the pyramids, there is an uncertainty of several degrees in the directions so determined. In short, a general and accurate method for the determination of electric axes of quartz has so far not yet been found.

Despite the remarkable progress made in the technique for testing the qualities of quartz crystals, a number of physical properties of quartz, notably the piezoelectric constant, the modulus of elasticity, the conductivities of heat and electricity are still found to vary from specimen to specimen in quantities definitely too great to be ascribed to experimental errors. For the explanation of such discrepancies, certain authors postulate the existence of undetectable microscopic twinnings in quartz; while others tend to believe that the crystalline network might show a number of structural faults of which the nature and the frequency of occurrence are functions of both the conditions of crystallization and the impurities present. In the present state of affairs, it is impossible to decide which of the above assumptions is more near to the truth. However, there is one thing which seems to be certain, i.e., the observed discrepancies are due to some unknown crystalline defects.

Since quartz is an anisotropic substance, any effect produced in it by an external agent will generally be directional in nature and may serve for the investigation of certain properties of the crystal. On this consideration, the writer studied the effect of electric field on the etching of quartz. New etching figures have thus been obtained. Hence new methods for the determination of electric axes and their polarities on the one hand, and new kinds of crystalline defects on the other, have been found.

EXPERIMENTAL

The principal apparatus used in the present investigations are two types of etching cells, a high tension transformer with some accessories for the production of d.c. voltage, a Norrenberg's polariscope for the determination of optic axis of quartz and a high precision micrometer for the determination of the orientation and the dimensions of the etching figures. One of the types of etching cell is made of a short pitch-coated glass tube with a quartz plate intended for study as the bottom. The inner surface of the plate is in contact with hydrofluoric acid while the outer one is coated with tinfoil so that an electric field normal to the plate may be set up in it. The second type of cell is in the form of V (see Fig. 1). It is made by cementing a glass tubing on each surface of a quartz plate, so that both surfaces of the plate may be etched at the same time. The quartz plate under investigation is subjected to the actions of the corrosive acid and the electric field simultaneously. The plates used are all cut perpendicular to the optic axis. They have a thickness varying from 1 to 2 millimeters. The surfaces of plates to be etched by acid are smoothed or polished. Both a.c. and d.c. fields have been used in the experiments. The highest voltage which the transformer can supply is about 40 kv. Owing to the lack of adequate voltmeter for measuring high tensions, the high voltages indicated in the article are values estimated from spark-gap lengths.

The experimental arrangement with the use of d.c. voltage is schematically shown in Fig. 1. In the figure, E represents the V-form etching cell. A "spark" rectifier, P, consists of a needle point

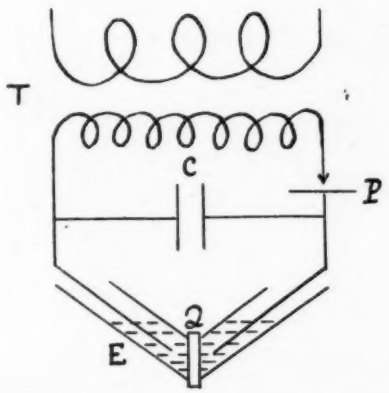


FIG. 1. E: etching cell, Q: quartz plate, C: condenser, P: "spark" rectifier.

[7] Choong Shin-piaw, Nature **154**, 464 (Oct. 7, 1944).

and a plate electrode connected in series with the cell and the secondary coil of the transformer, T. A condenser, C, of capacity some 0.005 mf is connected in parallel with the cell for stabilizing the d.c. voltage produced. In the case where an a.c. voltage is used, the rectifier, P, as well as the condenser, C, are removed and the two electrodes of the etching cell are then connected directly to the secondary coil of the transformer.

RESULTS

When a smoothed or polished surface of a quartz plate cut normal to the optic axis is in contact with hydrofluoric acid for some time, it becomes completely mat. This is due to the well-known fact that the etched surface becomes roughened by the formation of a great number of very small projecting triangular pyramids. Since the linear dimension of these pyramids rarely exceeds 0.02 mm, they can be observed only under a microscope. Now the writer finds that in case an electric field is applied normally to the surface of a plate during etching, this surface becomes mat as usual, but, in addition, an etching pattern of entirely different nature is formed. The elementary triangular or starry figures of which the new pattern is composed are usually large enough to be visible by the naked eye. The general aspects of the pattern as well as the micrographic structure of its elementary units change markedly with the strength of the applied field. On the other hand, the appearance and the orientation of the etching figures are independent of the signs of electrodes. The d.c. and a.c. fields are equally good for the production of the new figures. By the way, it may be of interest to point out here that the structure as well as the orientations of the projecting pyramids of corrosion which coexist with the new etching pattern, are not affected even in the slightest by the electric field.

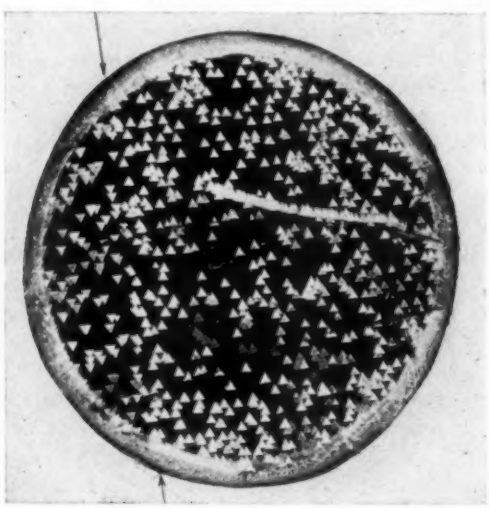


FIG. 2. Etching figures produced under weak field. Note the change of orientations of the triangular etching figures from one twinned part to another (dividing line indicated by arrows) and the crowding up of these figures along a polishing scratch.

I. Etching Pattern Caused by Weak Electric Field

By weak field is meant a field of intensity weaker than some 10,000 volts per millimeter, a superior limit assigned rather arbitrarily. The weakest field actually employed is about 3000 volts per millimeter. With this field and with a moderately strong acid (about 20 percent acid), the optimum duration of etching is around 40 hours, a duration some ten times longer than in the case with strong fields. Owing to the necessity of greatly prolonging the time of etching, the use of very weak fields is not practical.

The etching pattern produced under weak fields consists of a number of regularly orientated triangular figures. These figures are randomly distributed in most etched surfaces, but cases have been found where the figures show a tendency to form up in well-defined straight lines in certain localities. It is interesting to point out in this connection that the etching figures show a marked preference to develop along polishing scratches and surface cleavages. The close relation between the orientation of the etching figures and the atomic pattern of quartz is clearly manifested by the fact that the figures change their orientation from one twinned part to another (Fig. 2).

A microscopic examination reveals that the triangular etching figures may be classified into two categories, namely, the one with a blank equilateral triangle in the center (category A), and the other without it (category B), as diagrammatically represented in Fig. 3 and Fig. 4.

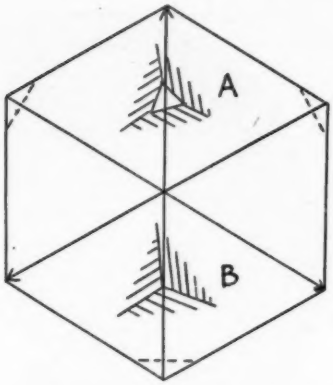


FIG. 3. Right-handed quartz.

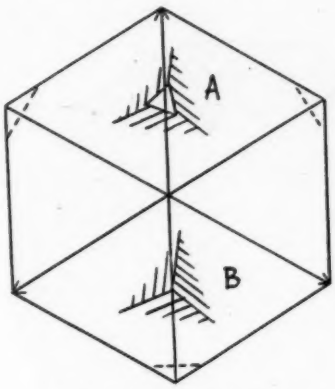


FIG. 4. Left-handed quartz.

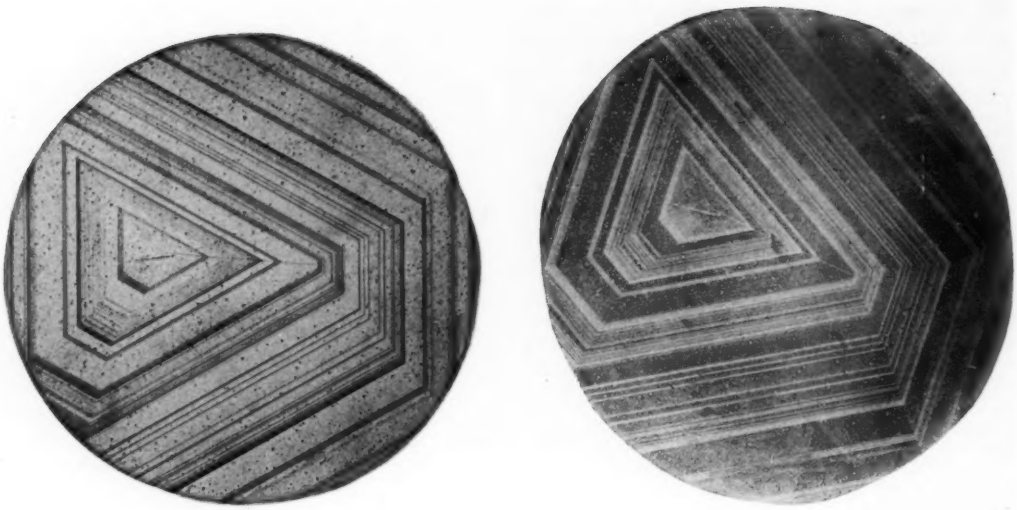
The regular hexagons of Fig. 3 and Fig. 4, represent the cross sections of the hexagonal prism of quartz. The triangular or starry figures *A* and *B* indicate the orientation and structure of the two categories of etching figures.

The figures of category A are of common occurrence while those of category B are rarely encountered except in the region of contact between the plate and the glass tubing. As may be seen from the diagrams, the etching figures are actually composed of three sets of parallel lines. There is a close relation between the way of arrangement of these sets of lines and the sense of optical rotation of quartz. For a right-handed quartz, the lines are found to be arranged as indicated in Fig. 3, and for a left-handed crystal, as indicated in Fig. 4. These sets of lines make 120° with each other. They are within a few degrees parallel to the three electric axes of quartz with their outward ends directed to the positive extremities of the axes by compression. Careful measurements show that the orientation of the etching figures is a function of the applied field (see Fig. 6). With respect to the electric axes, the figures of right-handed quartz are deviated in the anticlockwise sense and those of left-handed quartz in the clockwise sense.

II. Etching Pattern Caused by Strong Electric Field

The triangular etching figures produced under the action of a strong field are generally smaller in size but by far larger in number than in the case of weak field. Because of the denseness in figures, the boundary between the twinned parts of an etched surface is, in the present case, very neatly revealed. The most striking features exhibited in most of the etched surfaces under high tensions are the appearance of groups of parallel long lines and bands as well as a kind of well delimited etched zones which are similar in appearance to a twinned zone. These groups of parallel lines and bands change abruptly their directions at the boundaries of the new zones but never at those of the twinned zones. That such new zones are not of the same nature as the ordinary twinned zones is shown beyond any doubt by the fact that the former cannot be detected by optical means and ordinary etching process (i.e., etching not subject to field). The angles formed between the groups of parallel lines and bands are 120°, but occasionally some of the lines or bands may make 60° with each other (especially those lying in the central part of the hexagonal prism of the crystal). Careful comparisons made between the orientation of these groups of lines and the edges of a cross section of the hexagonal prism of quartz prove conclusively that these lines and bands are exactly parallel to the edges, i.e., parallel to the electric axes of quartz (Fig. 5).

The micrographic structure of many of the elementary starry figures, produced under strong field, is similar to that designated as figures of category B, indicated in Fig. 3 and Fig. 4. But there are many others whose structure is much simpler, for instead of exhibiting brushy lines, they possess only three main arms which are



(a) (b)

FIG. 5. Etching pattern produced under strong field. The groups of parrallel lines and bands as well as the new etched zones are clearly shown. A twinned zone is also observable in the plate. Note the parallelism between the natural edges of the plate and the corresponding groups of etching lines and bands. Figure 5a shows the etched plate by transmitted light while 5b shows it by reflected light.

concurrent at a point and make an angle of 120° with each other. The directions of the brushy lines or of the arms, as the case may be, do not coincide with those of the electric axes. As in the case of weak field, the deviation is anticlockwise for right-handed quartz and clockwise for a left-handed crystal. The elementary figures which compose the groups of lines and bands as well as the new etched zones are practically the same in structure and in orientation as those in the remaining parts of an etched surface. The only factor which renders the particular localities discernible to the naked eye is that more figures are developed in them. As has already been pointed out above, the polishing scratches and surface cleavages are also the seats of preference for the formation of elementary etching figures. Owing to the similarity in phenomena between these two cases, the writer is inclined to imagine that the presence of groups of lines and bands as well as the new etched zones is an evidence of the existence of "primary" crystal flaws. That ultramicroscopic flaws may exist in crystals is a well established fact since the works of A. Smekal[8] and others on salt crystals. According to Smekal, inclusions are the essential cause of the formation of flaws.

The new defects of quartz, namely, the so-called crystal flaws which are made conspicuous by the present method of experimentation are found very often in quartz crystals. In fact, more than 90 percent of the specimens under investigation show, more or less prominently, either the groups of long lines and bands or the new etched zones or both of them. Those who have the occasion to work a little quartz know that twinning is of frequent occurrence in quartz and that a block of perfectly homogeneous crystal is rare. To this, the writer may emphasize that it is much easier to find a piece of quartz free from twinning than from the new defects.

Some preliminary tests have been made to see whether the new defects of quartz will effect its oscillation characteristics. It is found that the plates made from a specimen which shows groups of prominent lines and bands are in general more difficult to set into vibration; moreover, they possess a resonance point less sharp. No doubt, from the technical point of view, it will be worth while to study how the new defects of quartz will

[8] A. Smekal, Physik. Zeits. **35**, p. 643 (1934); *International Conference on Physics* (1934), vol. 2 "Structure-Sensitive Properties of Salt Crystals."

influence the piezoelectric constant, the modulus of elasticity, the electric conductivity and many other properties of quartz.

It has been mentioned in several occasions that the orientation of the elementary etching figures varies with the intensity of applied field. In order to see in what manner this orientation varies with the intensity, a series of observations with different field strength was made in each of a number of plates. For the sake of simplicity, a.c. field was used in the investigation. As a typical example, the result obtained from one of the plates is represented in Fig. 6. In the figure, the ordinate stands for the angle of deviation, θ, i.e., angle between the main arms or the brushy lines of an etching figure and the corresponding electric axes of quartz. The abscissa instead of representing directly the field intensity, V_s, of which the exact value is unknown, represents the ratio (primary voltage/thickness of plate) $= V_p$. Of course, such a substitution requires the relation $V_s/V_p = C$ (constant) to hold. Since the current passing through a quartz plate under etching, i.e., passing through the secondary coil of the transformer, was negligibly small, the above relation must be approximately true. The constant, C, of the transformer used is roughly equal to 170. It may be inferred from the curve that as the field intensity, V_s, increases from small values, the angle of deviation, θ, increases with it, until a critical point (V_s', θ') is reached; since then, θ decreases and tends to a limit as V_s continues to increase. The part of the curve where V_s is smaller than V_s' may be represented fairly well by the expression,

$$\theta = K V_s,$$

where K is a constant of proportionality. The constant, K, and other characteristics of the curve are likely to change slightly from one specimen to another.

III. Methods for the Determination of Electric Axes of Quartz

The curve of Fig. 6 shows clearly that the orientation of the triangular or starry etching figures, produced under the action of an electric field, will, in no case, be the same as that of the electric axes of quartz. Nevertheless, it is possible, by means of weak field, to render the orientation of such figures as near to that of electric axes as one pleases. The only drawback to the use of very weak field is the necessity to prolong greatly the time of etching. As has been said before, most quartz crystals show, under the action of strong fields, groups of long etching lines and bands which are exactly parallel to the electric axes of quartz. Evidently these lines and bands may serve, most ideally, for the determination of quartz axes. What seems to be regrettable is that not all crystals will show such lines and bands; so that this method cannot be regarded as a general one. Owing to the fact that the orientation of the triangular figures, produced under strong fields, makes comparatively large angles with the electric axes of quartz, the former is not accurate enough to be taken as representative of the latter. However these very same etching figures may be used in an indirect way, for a precise determination of the axes in question. The procedure for such a determination is indicated below.

Since the triangular or starry etching figures produced under an electric field are deviated, with respect to the electric axes of quartz, either clockwise or anticlockwise according as the crystal is a left-handed or a right-handed specimen, it is clear that the figures, produced on both surfaces of a plate will appear to deviate in

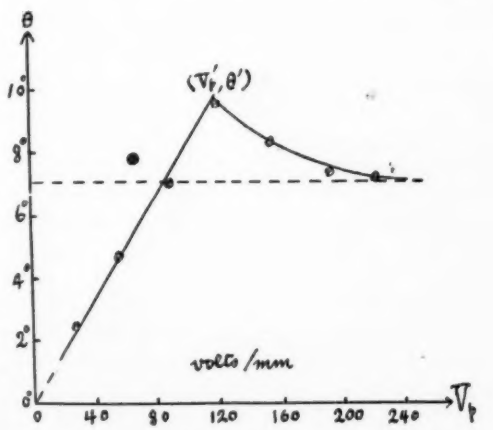


FIG. 6. The ordinate represents the angle of deviation, θ, of the triangular figures, and the abscissa the ration (primary voltage/plate thickness) $= V_p$. The actual field intensity, V_s, is roughly equal to $170 V_p$.

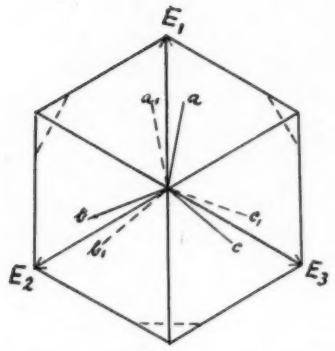


FIG. 7. a, b, and c represent the brushy lines or arms of a starry etching figure in the front surface, and a_1, b_1, and c_1 the brushy lines or arms of a figure in the back surface. E_1, E_2, and E_3 represent the direction of the electric axes of quartz.

opposite senses when viewed from any one surface. Now if the fields applied on both etched surfaces are of the same strength, the angle of deviation of figures in the two surfaces will be equal but in opposite senses as indicated in Fig. 7. Thus the bisectors of the three angles, aoa_1, bob_1 and coc_1 formed by the directions of the corresponding brushy lines or arms of etching figures in both surfaces coincide exactly with the directions of the three electric axes of quartz. By this method, the axes can be determined within the accuracy of one degree. As to the polarity of the axes, it is found that the outward ends of the bisectors represent the positive extremities of the axes by compression. Since the triangular or starry etching figures can always be produced in any specimen of quartz, the method is necessarily a general one.

The writer has much pleasure in expressing thanks to Dr. Li Shu-Hua, President of the National Academy of Peiping, and to Dr. Ny Tsi-Ze, Director of the Institute of Physics of the same Academy, for much helpful discussion. Thanks are also due Mr. Choong Shin-San for his help in the work.